CDS

TECHNICAL MEMORANDUM NO. CIT-CDS 96-020
November, 1996

"The Hamiltonian and Lagrangian Approaches to the Dynamics of Nonholonomic Systems"

Jerrold E. Marsden and Wang-Sang Koon

Control and Dynamical Systems
California Institute of Technology
Pasadena, California 91125

TECHNICAL MEMORANDUM: CIT/CDS 96-020

The Hamiltonian and Lagrangian Approaches to the Dynamics of Nonholonomic Systems

Wang-Sang Koon
Department of Mathematics
University of California
Berkeley, CA 94720-3840
koon@math.berkeley.edu

Jerrold E. Marsden*
Control and Dynamical Systems
California Institute of Technology 116-81
Pasadena, CA 91125
marsden@cds.caltech.edu

November 11, 1996

Abstract

This paper compares the Hamiltonian approach to systems with nonholonomic constraints (see Weber [1982], Arnold [1988], and Bates and Snyatycki [1992], Van der Schaft and Maschke [1994] and references therein) with the Lagrangian approach (see Koiller [1992], Ostrowski [1996] and Bloch, Krishnaprasad, Marsden and Murray [1996]). There are many differences in the approaches and each has its own advantages; some structures have been discovered on one side and their analogues on the other side are interesting to clarify. For example, the momentum equation and the reconstruction equation was first found on the Lagrangian side and is useful for the control theory of these systems, while the failure of the reduced two form to be closed (i.e., the failure of the Poisson bracket to satisfy the Jacobi identity) was first noticed on the Hamiltonian side. Clarifying the relation between these approaches is important for the future development of the control theory and stability and bifurcation theory for such systems. In addition to this work, we treat, in this unified framework, a simplified model of the bicycle (see Getz [1994] and Getz and Marsden [1995]), which is an important underactuated (nonminimum phase) control system.

Contents

*Research partially supported by the NSF, DOE, a Fairchild Fellowship at Caltech, and the Fields Institute for Research in the Mathematical Sciences.

1 Introduction

The General Setting. Many important problems in robotics, the dynamics of wheeled vehicles and motion generation, involve nonholonomic mechanics. Some of the important issues are trajectory tracking, dynamic stabilty and feedback stabilization (including nonminimum phase systems), bifurcation and control. Many of these systems have symmetry, such as the group of Euclidean motions in the plane or in space and this symmetry plays an important role in the theory.

In the last several years, several basic works have been done on both the Hamiltonian and the Lagrangian sides of the theory. Papers like Weber [1986], Koiller [1992], Bloch and Crouch [1992], Krishnaprasad, Dayawansa and Yang [1992, 1993], Bates and Sniatycki [1993], van der Schaft and Maschke [1994], Ostrowski [1996] and Bloch, Krishnaprasad, Marsden and Murray [1996] have laid a firm foundation for understanding nonholonomic mechanics with symmetry.

Bates and Sniatycki [1993], hereafter denoted [BS], developed the Hamiltonian side, while Bloch, Krishnaprasad, Marsden and Murray [1996], hereafter denoted [BKMM], has explored the Lagrangian side. Our aim is to establish links between these two sides and use the ideas and results of each to shed light on the other, with the goal of deepening our understanding of both points of view. We hope that it will aid related efforts such as extending the results of energy-momentum method for stability of relative equilibria and the theory of Hamiltonian bifurcations to nonholonomic mechanics. In the spirit of [BKMM], we do many of the calculations in coordinates to help in the study of examples.

We illustrate the basic theory with the snakeboard, the well known example treated in [BKMM]. We also treat a simplified model of the bicycle (introduced in Getz [1994] and Getz and Marsden [1995]). This is an important prototype control system because it is an underactuated balance system.

Outline of the Paper. We begin in §2 by recalling some of the main results of [BKMM] and of [BS] in the general context of nonholonomically constrained systems. In that section, we establishing the precise link between them. The snakeboard example is begun in this section.

In §3, we treat systems with symmetry and study the momentum equation, the reconstruction equation and the reduced Lagrange-d'Alembert equations from both the Hamiltonian and the Lagrangian points of view. This clarifies which construction in [BS] corresponds to the momentum equation of [BKMM]. This section also continues the snakeboard example and treats the bicycle.

Summary of the Main Results. The main results of the present work are as follows:

- The precise relation between the constructions in the papers [BS] and [BKMM] are given.

- The reduced Lagrange-d'Alembert equations established in [BKMM] are shown to be equivalent to the reduced nonholonomic Hamilton equations implicitly given in [BS].
- The relation between the constructions is illustrated with two specific examples, the snakeboard and the nonholonomically constrained particle.
- A simplified model of the bicycle is treated.

2 General Nonholonomic Mechanical Systems

Following the approaches of both [BS] and [BKMM], we first consider mechanics in the presence of homogeneous linear nonholonomic velocity constraints. For now, no symmetry assumptions are made; we add such assumptions in the following sections.

In this section,

1. we recall the basic ideas and results from [BKMM] on general nonholonomic systems: in particular, how to describe constraints using Ehresmann connections and how to write the Lagrange d'Alembert equations of motion using the curvature of this connection.
2. We review the geometric structure of Hamiltonian systems with nonholonomic constraints from [BS], including a general procedure for finding the equations of motion for nonholonomic systems from the Hamiltonian point of view.
3. We construct the geometric objects on the Lagrangian side corresponding to those on the Hamiltonian side using the Legendre transformation in the context of nonholonomic constraints.
4. We prove that these dual procedures gives us the same Lagrange d'Alembert equations as in [BKMM]. Since this proof is done in coordinates, it also provides a concrete coordinate based procedure for finding the equations of motion on the Hamiltonian side.
5. We will use the Hamiltonian procedure to work out the example of snakeboard taken from [BKMM].

2.1 Review of the Lagrangian Approach

We start with a configuration space Q with local coordinates denoted $q^i, i = 1, \ldots, n$ and a distribution $\mathcal{D}$ on Q that describes the kinematic nonholonomic constraints. The distribution is given by the specification of a linear subspace $\mathcal{D}_q \subset T_qQ$ of the tangent space to Q at each point $q \in Q$.

In this paper we consider only homogeneous velocity constraints. The extension to affine constraints is straightforward, as in [BKMM]. The extension to constraints nonlinear in the velocity is also not difficult, but it requires one to work on a higher tangent bundle, which we do not describe in this paper.

The dynamics of a nonholonomically constrained mechanical system is governed by the Lagrange d'Alembert principle. The principle states that the equations of motion of a curve $q(t)$ in configuration space are obtained by setting to zero the variations in the integral of the Lagrangian subject to variations lying in the constraint distribution and that the velocity of the curve $q(t)$ itself satisfies the constraints; that is, $\dot{q}(t) \in \mathcal{D}_{q(t)}$. Standard arguments in the calculus of variations show that this "constrained variational principle" is equivalent to the equations

$$-\delta L := \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} \right) \delta q^i = 0, \tag{2.1.1}$$

for all variations δq such that $\delta q \in \mathcal{D}_q$ at each point of the underlying curve $q(t)$. These equations are often equivalently written as

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = \lambda_i, \tag{2.1.2}$$

where λ_i is a set of Lagrange multipliers ($i = 1, \ldots, n$), representing the force of constraint. Intrinsically, this multiplier λ is a section of the cotangent bundle over $q(t)$ that annihilates the constraint distribution. The Lagrange multipliers are often determined by using the condition that $\dot{q}(t)$ lies in the distribution.

In Bloch and Crouch [1992] and Lewis [1996], the Lagrange d'Alembert equations are shown to have the form of a generalized acceleration condition

$$\nabla_{\dot{q}}\dot{q} = 0$$

for a suitable affine connection on Q and the force of constraint λ is interpreted as a generalized second fundamental form (as is well known for systems with holonomic constraints; see Abraham and Marsden [1978], for example). In this form of the equations, one can add external forces directly to the right hand sides so that the equations now become in the form of a generalized Newton law. This form is convenient for control purposes.

To explore the structure of the Lagrange d'Alembert equations in more detail, let $\{\omega^a\}$, $a = 1, \ldots, k$ be a set of k independent one forms whose vanishing describes the constraints; i.e., the distribution $\mathcal{D}$. One can introduce local coordinates $q^i = (r^\alpha, s^a)$ where $\alpha = 1, \ldots n - k$, in which ω^a has the form

$$\omega^a(q) = ds^a + A^a_\alpha(r, s)dr^\alpha \tag{2.1.3}$$

where the summation convention is in force. In other words, we are locally writing the distribution as

$$\mathcal{D} = \{(r, s, \dot{r}, \dot{s}) \in TQ \mid \dot{s} + A^a_\alpha \dot{r}^\alpha = 0\}.$$

The equations of motion, (2.1.1) may be rewritten by noting that the allowed variations $\delta q^i = (\delta r^\alpha, \delta s^a)$ satisfy $\delta s^a + A^a_\alpha \delta r^\alpha = 0$. Substitution into (2.1.1) gives

$$\left(\frac{d}{dt}\frac{\partial L}{\partial \dot{r}^\alpha} - \frac{\partial L}{\partial r^\alpha}\right) = A^a_\alpha \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{s}^a} - \frac{\partial L}{\partial s^a}\right). \tag{2.1.4}$$

Equation (2.1.4) combined with the constraint equations

$$\dot{s}^a = -A^a_\alpha \dot{r}^\alpha \tag{2.1.5}$$

gives a complete description of the equations of motion of the system; this procedure may be viewed as one way of eliminating the Lagrange multipliers. Using this notation, one finds that $\lambda = \lambda_a \omega^a$, where $\lambda_a = \frac{d}{dt}\frac{\partial L}{\partial \dot{s}^a} - \frac{\partial L}{\partial s^a}$.

Equations (2.1.4) can be written in the following way:

$$\frac{d}{dt}\frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A^a_\alpha \frac{\partial L_c}{\partial s^a} = -\frac{\partial L}{\partial \dot{s}^b} B^b_{\alpha\beta}\dot{r}^\beta, \tag{2.1.6}$$

where

$$L_c(r^\alpha, s^a, \dot{r}^\alpha) = L(r^\alpha, s^a, \dot{r}^\alpha, -A^a_\alpha(r, s)\dot{r}^\alpha).$$

is the coordinate expression of the constrained Lagrangian defined by $L_c = L|\mathcal{D}$ and where

$$B^b_{\alpha\beta} = \left(\frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} + A^a_\alpha \frac{\partial A^b_\beta}{\partial s^a} - A^a_\beta \frac{\partial A^b_\alpha}{\partial s^a}\right). \tag{2.1.7}$$

Letting $d\omega^b$ be the exterior derivative of ω^b, a computation shows that

$$d\omega^b(\dot{q}, \cdot) = B^b_{\alpha\beta}\dot{r}^\alpha dr^\beta$$

and hence the equations of motion have the form

$$-\delta L_c = \left(\frac{d}{dt}\frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A^a_\alpha \frac{\partial L_c}{\partial s^a} \right) \delta r^\alpha = -\frac{\partial L}{\partial \dot{s}^b} d\omega^b(\dot{q}, \delta r).$$

This form of the equations isolates the effects of the constraints, and shows, in particular, that in the case where the constraints are integrable (*i.e.,* $\mathbf{d}\omega = 0$), the equations of motion are obtained by substituting the constraints into the Lagrangian and then setting the variation of L_c to zero. However in the non-integrable case the constraints generate extra (curvature) terms, which must be taken into account.

The above coordinate results can be put into an interesting and useful intrinsic geometric framework. The intrinsically given information is the distribution and the Lagrangian. *Assume* that there is a bundle structure $\pi_{Q,R} : Q \to R$ for our space Q, where R is the *base* manifold and $\pi_{Q,R}$ is a submersion and the kernel of $T_q\pi_{Q,R}$ at any point $q \in Q$ is called the *vertical space* V_q. One can always do this locally. An *Ehresmann connection* A is a vertical valued one form on Q such that

1. $A_q : T_qQ \to V_q$ is a linear map and
2. A is a projection: $A(v_q) = v_q$ for all $v_q \in V_q$.

Hence, $T_qQ = V_q \oplus H_q$ where $H_q = \ker A_q$ is the *horizontal space at* q, sometimes denoted hor_q. Thus, an Ehresmann connection gives us a way to split the tangent space to Q at each point into a horizontal and vertical part.

If the Ehresmann connection is chosen in such a way that the given constraint distribution $\mathcal{D}$ is the horizontal space of the connection; that is, $H_q = \mathcal{D}_q$, then in the bundle coordinates $q^i = (r^\alpha, s^a)$, the map $\pi_{Q,R}$ is just projection onto the factor r and the connection A can be represented locally by a vector valued differential form ω^a:

$$A = \omega^a \frac{\partial}{\partial s^a}, \qquad \omega^a(q) = ds^a + A^a_\alpha(r,s)dr^\alpha,$$

and the horizontal projection is the map

$$(\dot{r}^\alpha, \dot{s}^a) \mapsto (\dot{r}^\alpha, -A^a_\alpha(r,s)\dot{r}^\alpha). \tag{2.1.8}$$

The curvature of an Ehresmann connection A is the vertical valued two form defined by its action on two vector fields X and Y on Q as

$$B(X,Y) = -A([\mathrm{hor}\, X, \mathrm{hor}\, Y])$$

where the bracket on the right hand side is the Jacobi-Lie bracket of vector fields obtained by extending the stated vectors to vector fields. This definition shows the sense in which the curvature measures the failure of the constraint distribution to be integrable.

In coordinates, one can evaluate the curvature B of the connection A by the following formula:

$$B(X,Y) = \mathbf{d}\omega^a(\mathrm{hor}\, X, \mathrm{hor}\, Y)\frac{\partial}{\partial s^a},$$

so that the local expression for curvature is given by

$$B(X,Y)^a = B^a_{\alpha\beta}X^\alpha Y^\beta \tag{2.1.9}$$

where the coefficients $B^a_{\alpha\beta}$ are given by (2.1.7).

The Lagrange d'Alembert equations may be written intrinsically as

$$\delta L_c = \langle \mathbb{F}L, B(\dot{q}, \delta q)\rangle,$$

in which δq is a horizontal variation (*i.e.,* it takes values in the horizontal space) and B is the curvature regarded as a vertical valued two form, in addition to the constraint equations

$$A(q) \cdot \dot{q} = 0.$$

Here $\langle\, ,\rangle$ denotes the pairing between a vector and a dual vector and

$$\delta L_c = \left\langle \delta r^\alpha, \frac{\partial L_c}{\partial r^\alpha} - \frac{d}{dt}\frac{\partial L_c}{\partial \dot{r}^\alpha} - A^a_\alpha \frac{\partial L_c}{\partial s^a} \right\rangle.$$

As shown in [BKMM], when there is a symmetry group G present, there is a natural bundle one can work with and put a connection on, namely the bundle $Q \to Q/G$. In the generality of the preceding discussion, one can get away with just the distribution itself and can introduce the corresponding Ehresmann connection locally. In fact, the bundle structure $Q \to R$ is really a "red Herring". The notion of curvature as a $T_qQ/\mathcal{D}_q$ valued form makes good sense and is given locally by the same expressions as above. However, keeping in mind that we eventually want to deal with symmetries and in that case there is a natural bundle, the Ehresmann assumption is nevertheless a reasonable bridge to the more interesting case with symmetries.

2.2 Review of the Hamiltonian Formulation

The approach of [BS] starts on the Lagrangian side with a configuration space Q and a Lagrangian L of the form kinetic energy minus potential energy, *i.e.,*

$$L(q, \dot{q}) = \frac{1}{2}\langle\langle \dot{q}, \dot{q}\rangle\rangle - V(q),$$

where $\langle\langle\; ,\; \rangle\rangle$ is a metric on Q defining the kinetic energy and V is a potential energy function. We do not restrict ourselves to Lagrangians of this form.

As above, our nonholonomic constraints are given by a distribution $\mathcal{D} \subset TQ$. We also let $\mathcal{D}^{\mathrm{o}} \subset T^*$ Q denote the annihilator of this distribution.

As above, the basic equations are given by the Lagrange d'Alembert principle.

The Legendre transformation $\mathbb{F}L : TQ \to T^*Q$, assuming that it is a diffeomorphism, is used to define the Hamiltonian $H : T^*Q \to \mathbb{R}$ in the standard fashion (ignoring the constraints for the moment):

$$H = \langle p, \dot{q}\rangle - L = p_i\dot{q}^i - L.$$

Here, the momentum is $p = \mathbb{F}L(v_q) = \partial L/\partial \dot{q}$. Under this change of variables, the equations of motion are written in the Hamiltonian form as

$$\begin{aligned} \dot{q}^i &= \frac{\partial H}{\partial p_i}. \\ \dot{p}_i &= -\frac{\partial H}{\partial q^i} + \lambda_a \omega^a_i, \end{aligned}$$

where $i = 1, \ldots, n$, together with the constraint equations.

The preceding Hamilton equations can be rewritten as

$$X \,\lrcorner\, \Omega = dH + \lambda_a \pi_Q^* \omega^a, \tag{2.2.1}$$

where X is the vector field on T^*Q governing the dynamics, Ω is the canonical symplectic form on T^*Q, and $\pi_Q : T^*Q \to Q$ is the cotangent bundle projection. We may write X in coordinates as $X = \dot{q}^i \partial_{q^i} + \dot{p}_i \partial_{p_i}$.

On Lagrangian side, we saw that one can get rid of the Lagrangian multipliers. On the Hamiltonian side, it is also desireable to model the Hamiltonian equations without the Lagrange multipliers by a vector field on a submanifold of T^*Q. We do this in what follows.

First of all, we define the set $\mathcal{M} = \mathbb{F}L(\mathcal{D}) \subset T^*Q$, so that the constraints on the Hamiltonian side are given by

$$p \in \mathcal{M}.$$

Besides $\mathcal{M}$, another basic object we deal with is defined as

$$\mathcal{F} = (T\pi_Q)^{-1}(\mathcal{D}) \subset TT^*Q.$$

Using a basis ω^a of the annihilator $\mathcal{D}^o$, we can write these spaces as

$$\mathcal{M} = \{p \in T^*Q \mid \omega^a((\mathbb{F}L)^{-1}(p)) = 0\}, \tag{2.2.2}$$

and

$$\mathcal{F} = \{u \in TT^*Q \mid \langle \pi_Q^* \omega^a, u \rangle = 0\}. \tag{2.2.3}$$

Finally, we define

$$\mathcal{H} = \mathcal{F} \cap T\mathcal{M}.$$

Using natural coordinates $(q^i, p_i, \dot{q}^i, \dot{p}_i)$ on TT^*Q, we see that the distribution $\mathcal{F}$ naturally lifts the constraint on $\dot{q}$ from TQ to TT^*Q. On the other hand, the space $\mathcal{M}$ puts the associated constraints on the variable p and therefore the intersection $\mathcal{H}$ puts the constraints on both variables.

To eliminate the Lagrange multipliers, we regard the Hamiltonian equations as a vector field on the constraint submanifold $\mathcal{M} \subset T^*Q$ which takes values in the constraint distribution $\mathcal{H}$. Next we recall from [BS] how to construct these equations intrinsically using the ideas of symplectic geometry.

A result of [BS] is that $\Omega_{\mathcal{H}}$, the restriction of the canonical two-form Ω of T^*Q fiberwise to the distribution $\mathcal{H}$ of the constraint submanifold $\mathcal{M}$, is nondegenerate. Note that $\Omega_{\mathcal{H}}$ is not a true two form on a manifold, so it does not make sense to speak about it being closed. We speak of it as a fiber-restricted two form to avoid any confusion. Of course it still makes sense to talk about it being nondegenerate; it just means nondegenerate as a bilinear form on each fiber of $\mathcal{H}$. The dynamics is then given by the vector field $X_{\mathcal{H}}$ on $\mathcal{M}$ which takes values in the constraint distribution $\mathcal{H}$ and is determined by the condition

$$X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{H}} = dH_{\mathcal{H}} \tag{2.2.4}$$

where $dH_{\mathcal{H}}$ is the restriction of $dH_{\mathcal{M}}$ to $\mathcal{H}$. We will be exploring the coordinate meaning of this condition and its comparison with the Lagrangian formulation in the subsequent sections.

2.3 Lagrangian Side

We now construct the geometric structures on the tangent bundle TQ corresponding to those on the Hamiltonian side from the preceding subsection and formulate a similar procedure for obtaining the equations of motion. By doing this, it will be easier to made comparision with the geometric constructions and analytic formulations in [BKMM].

First of all, we can define the energy function E simply as $E = H \circ \mathbb{F}L$ and pull back to TQ the canonical two-form on T^*Q and denote it by Ω_L.

We define the distribution

$$\mathcal{C} = (T\tau_Q)^{-1}(\mathcal{D}) \subset TTQ,$$

where $\tau_Q : TQ \to Q$. In coordinates, the distribution $\mathcal{C}$ consists of vectors annihilated by the form $\tau_Q^* \omega^a$:

$$\mathcal{C} = \{u \in TTQ \mid \langle \tau_Q^* \omega^a, u \rangle = 0\}. \tag{2.3.1}$$

When $\mathcal{C}$ is restricted to the constraint submanifold $\mathcal{D} \subset TQ$, we obtain the constraint distribution $\mathcal{K}$:

$$\mathcal{K} = \mathcal{C} \cap T\mathcal{D}. \tag{2.3.2}$$

Clearly $\mathcal{M} = \mathbb{F}L(\mathcal{D})$ and $\mathcal{H} = T\mathbb{F}L(\mathcal{K})$.

The dynamics is given by a vector field $X_{\mathcal{K}}$ on the manifold $\mathcal{D}$ which takes values in $\mathcal{K}$ and satisfies the equation

$$X_{\mathcal{K}} \lrcorner\, \Omega_{\mathcal{K}} = dE_{\mathcal{K}}, \tag{2.3.3}$$

where $dE_{\mathcal{K}}$ and $\Omega_{\mathcal{K}}$ are the restrictions of $dE_{\mathcal{D}}$ and $\Omega_{\mathcal{D}}$ respectively to the distribution $\mathcal{K}$ and where $E_{\mathcal{D}}$ and $\Omega_{\mathcal{D}}$ are the restrictions of E and Ω_L to $\mathcal{D}$.

2.4 The equivalence of the Hamiltonian and the Lagrange-d'Alembert formulations

The Lagrangian procedure on TQ formulated in the preceding subsection act as a bridge between [BS] and [BKMM]. We can show the correctness of the Lagrangian procedure given above by (carefully) invoking the results of [BS] (generalized to arbitrary Lagrangians and with some gaps filled in), or by checking the methods against the results of [BKMM]. We choose the latter method.

Theorem 2.1 *Consider a configuration space Q, a hyperregular Lagrangian L and a distribution $\mathcal{D}$ that describes the kinematic nonholonomic constraints. The $\mathcal{K}$-valued vector field $X_{\mathcal{K}}$ on $\mathcal{D}$ given by the equation*

$$X_{\mathcal{K}} \lrcorner\, \Omega_{\mathcal{K}} = dE_{\mathcal{K}} \tag{2.4.1}$$

defines dynamics that are equivalent to the Lagrange d'Alembert equations together with the constraints.

Proof To keep things concrete and to provide additional insight, we shall give a coordinate based proof. Introduce local coordinates $(r^\alpha, s^a, \dot{r}^\alpha, \dot{s}^a)$ for TQ as described earlier. Local coordinates for the manifold $\mathcal{D}$ are given by $(r^\alpha, s^a, \dot{r}^\alpha)$.

Let us first compute $dE_{\mathcal{D}}$ and $X_{\mathcal{K}} \lrcorner\, \Omega_{\mathcal{D}}$. We claim that

$$E_{\mathcal{D}} = \frac{\partial L_c}{\partial \dot{r}^\beta} \dot{r}^\beta - L_c,$$

where $L_c = L|\mathcal{D}$ is the constrained Lagrangian. This is because $E = \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i - L$ and so restricting it to $\mathcal{D}$ we get

$$\begin{aligned} E_{\mathcal{D}} &= \left(\frac{\partial L}{\partial \dot{r}^\alpha} \dot{r}^\alpha + \frac{\partial L}{\partial \dot{s}^a} \dot{s}^a \right)\bigg|_{\mathcal{D}} - L_c \\ &= \frac{\partial L_c}{\partial \dot{r}^\alpha} \dot{r}^\alpha + A^a_\alpha \frac{\partial L}{\partial \dot{s}^a} \dot{r}^\alpha - A^a_\alpha \frac{\partial L}{\partial \dot{s}^a} \dot{r}^\alpha - L_c \\ &= \frac{\partial L_c}{\partial \dot{r}^\alpha} \dot{r}^\alpha - L_c. \end{aligned}$$

The differential of $E_{\mathcal{D}}$ is then computed to be

$$\begin{aligned} dE_{\mathcal{D}} &= \dot{r}^\beta \frac{\partial^2 L_c}{\partial r^\alpha \partial \dot{r}^\beta} dr^\alpha + \dot{r}^\beta \frac{\partial^2 L_c}{\partial \dot{r}^\alpha \partial \dot{r}^\beta} d\dot{r}^\alpha + \dot{r}^\beta \frac{\partial^2 L_c}{\partial s^b \partial \dot{r}^\beta} ds^b - \frac{\partial L_c}{\partial r^\alpha} dr^\alpha - \frac{\partial L_c}{\partial s^b} ds^b \\ &= \dot{r}^\beta \frac{\partial^2 L_c}{\partial r^\alpha \partial \dot{r}^\beta} dr^\alpha + \dot{r}^\beta \frac{\partial^2 L_c}{\partial \dot{r}^\alpha \partial \dot{r}^\beta} d\dot{r}^\alpha - \dot{r}^\beta A^b_\alpha \frac{\partial^2 L_c}{\partial s^b \partial \dot{r}^\beta} dr^\alpha + \dot{r}^\beta \frac{\partial^2 L_c}{\partial s^b \partial \dot{r}^\beta} (ds^b + A^b_\alpha dr^\alpha) \\ &\quad - \frac{\partial L_c}{\partial r^\alpha} dr^\alpha + A^b_\alpha \frac{\partial L_c}{\partial s^b} dr^\alpha - \frac{\partial L_c}{\partial s^b} (ds^b + A^b_\alpha dr^\alpha) \end{aligned}$$

As for $\Omega_{\mathcal{D}}$, we have

$$
\begin{aligned}
\Omega_{\mathcal{D}} &= -d\left(\left.\frac{\partial L}{\partial \dot{q}^i}\right|_{\mathcal{D}} dq^i\right) \\
&= -d\left\{\left(\frac{\partial L_c}{\partial \dot{r}^\beta} + A^b_\beta \frac{\partial L}{\partial \dot{s}^b}\right) dr^\beta + \frac{\partial L}{\partial \dot{s}^b} ds^b\right\} \\
&= -d\left(\frac{\partial L_c}{\partial \dot{r}^\beta}\right) \wedge dr^\beta - d\left(A^b_\beta \frac{\partial L}{\partial \dot{s}^b}\right) \wedge dr^\beta - d\left(\frac{\partial L}{\partial \dot{s}^b}\right) \wedge ds^b \\
&= -d\left(\frac{\partial L_c}{\partial \dot{r}^\beta}\right) \wedge dr^\beta - d\left(A^b_\beta \frac{\partial L}{\partial \dot{s}^b}\right) \wedge dr^\beta + A^b_\beta d\left(\frac{\partial L}{\partial \dot{s}^b}\right) \wedge dr^\beta - d\left(\frac{\partial L}{\partial \dot{s}^b}\right) \wedge (ds^b + A^b_\beta dr^\beta) \\
&= -d\left(\frac{\partial L_c}{\partial \dot{r}^\beta}\right) \wedge dr^\beta - \frac{\partial L}{\partial \dot{s}^b} d(A^b_\beta) \wedge dr^\beta - d\left(\frac{\partial L}{\partial \dot{s}^b}\right) \wedge (ds^b + A^b_\beta dr^\beta) \\
&= -\frac{\partial^2 L_c}{\partial r^\alpha \partial \dot{r}^\beta} dr^\alpha \wedge dr^\beta - \frac{\partial^2 L_c}{\partial \dot{r}^\alpha \partial \dot{r}^\beta} d\dot{r}^\alpha \wedge dr^\beta - \frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^\beta} ds^a \wedge dr^\beta \\
&\quad -\frac{\partial L}{\partial \dot{s}^b}\frac{\partial A^b_\beta}{\partial r^\alpha} dr^\alpha \wedge dr^\beta - \frac{\partial L}{\partial \dot{s}^b}\frac{\partial A^b_\beta}{\partial s^a} ds^a \wedge dr^\beta - d\left(\frac{\partial L}{\partial \dot{s}^b}\right) \wedge (ds^b + A^b_\beta dr^\beta) \\
&= \left(-\frac{\partial^2 L_c}{\partial r^\alpha \partial \dot{r}^\beta} + A^a_\alpha \frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^\beta} - \frac{\partial L}{\partial \dot{s}^b}\frac{\partial A^b_\beta}{\partial r^\alpha} + A^a_\alpha \frac{\partial L}{\partial \dot{s}^b}\frac{\partial A^b_\beta}{\partial s^a}\right) dr^\alpha \wedge dr^\beta \\
&\quad -\frac{\partial^2 L_c}{\partial \dot{r}^\alpha \partial \dot{r}^\beta} d\dot{r}^\alpha \wedge dr^\beta - \frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^\beta}(ds^a + A^a_\alpha dr^\alpha) \wedge dr^\beta \\
&\quad -\frac{\partial L}{\partial \dot{s}^b}\frac{\partial A^b_\beta}{\partial s^a}(ds^a + A^a_\alpha dr^\alpha) \wedge dr^\beta - d\left(\frac{\partial L}{\partial \dot{s}^b}\right) \wedge (ds^b + A^b_\beta dr^\beta),
\end{aligned}
$$

where there is a sum on all α, β.

Now we are ready to find the equations of motion. Any vector field $X_{\mathcal{D}}$ on $\mathcal{D}$ has the following coordinate form:

$$X_{\mathcal{D}} = \dot{r}^\alpha \partial_{r^\alpha} + \dot{s}^a \partial_{s^a} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha}.$$

Since $X_{\mathcal{K}}$ lies in the distribution $\mathcal{K}$, it is annihilated by the one-forms $ds^a + A^a_\alpha dr^\alpha$ and hence must be of the form

$$X_{\mathcal{K}} = \dot{r}^\alpha \partial_{r^\alpha} - A^a_\alpha \dot{r}^\alpha \partial_{s^a} + \ddot{r}^\alpha \partial \dot{r}^\alpha,$$

i.e., for the vector field $X_{\mathcal{K}}$,

$$\dot{s}^a = -A^a_\alpha \dot{r}^\alpha.$$

As for $X_{\mathcal{K}} \lrcorner\, \Omega_{\mathcal{K}}$, we have

$$
\begin{aligned}
X_{\mathcal{K}} \lrcorner\, \Omega_{\mathcal{K}} &= \dot{r}^\alpha \left(-\frac{\partial^2 L_c}{\partial r^\alpha \partial \dot{r}^\beta} + A^a_\alpha \frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^\beta} - \frac{\partial L}{\partial \dot{s}^b}\frac{\partial A^b_\beta}{\partial r^\alpha} + A^a_\alpha \frac{\partial L}{\partial \dot{s}^b}\frac{\partial A^b_\beta}{\partial s^a}\right) dr^\beta \\
&\quad - \dot{r}^\beta \left(-\frac{\partial^2 L_c}{\partial r^\alpha \partial \dot{r}^\beta} + A^a_\alpha \frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^\beta} - \frac{\partial L}{\partial \dot{s}^b}\frac{\partial A^b_\beta}{\partial r^\alpha} + A^a_\alpha \frac{\partial L}{\partial \dot{s}^b}\frac{\partial A^b_\beta}{\partial s^a}\right) dr^\alpha \\
&\quad + \dot{r}^\beta \frac{\partial^2 L_c}{\partial \dot{r}^\alpha \partial \dot{r}^\beta} d\dot{r}^\alpha - \ddot{r}^\alpha \frac{\partial^2 L_c}{\partial \dot{r}^\alpha \partial \dot{r}^\beta} dr^\beta
\end{aligned}
$$

This is because $X_{\mathcal{K}} \lrcorner\, \Omega_{\mathcal{K}}$ is the restriction of $X_{\mathcal{K}} \lrcorner\, \Omega_{\mathcal{D}}$ to the distribution of $\mathcal{K}$ and hence all the terms in $\Omega_{\mathcal{D}}$ which involve $(ds^b + A^b_\beta dr^\beta)$ vanish. The same is true for $dE_{\mathcal{D}}$.

Equating $X_{\mathcal{K}} \lrcorner\, \Omega_{\mathcal{K}}$ with $dE_{\mathcal{K}}$ and recalling that we have already obtained $\dot{s}^a = -A^a_\alpha \dot{r}^\alpha$, we get the following set of equations

$$\dot{r}^\beta \frac{\partial^2 L_c}{\partial r^\alpha \partial \dot{r}^\beta} - \dot{r}^\beta A^b_\alpha \frac{\partial^2 L_c}{\partial s^b \partial \dot{r}^\beta} - \frac{\partial L_c}{\partial r^\alpha} + A^b_\alpha \frac{\partial L_c}{\partial s^b}$$

$$
\begin{aligned}
= & \ \dot{r}^{\beta}\left(-\frac{\partial^2 L_c}{\partial r^{\beta}\partial \dot{r}^{\alpha}} + A^a_{\beta}\frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^{\alpha}} - \frac{\partial L}{\partial \dot{s}^b}\frac{\partial A^b_{\alpha}}{\partial r^{\beta}} + A^a_{\beta}\frac{\partial L}{\partial \dot{s}^b}\frac{\partial A^b_{\alpha}}{\partial s^a}\right) \\
& -\dot{r}^{\beta}\left(-\frac{\partial^2 L_c}{\partial r^{\alpha}\partial \dot{r}^{\beta}} + A^a_{\alpha}\frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^{\beta}} - \frac{\partial L}{\partial \dot{s}^b}\frac{\partial A^b_{\beta}}{\partial r^{\alpha}} + A^a_{\alpha}\frac{\partial L}{\partial \dot{s}^b}\frac{\partial A^b_{\beta}}{\partial s^a}\right) - \ddot{r}^{\beta}\frac{\partial^2 L_c}{\partial \dot{r}^{\beta}\partial \dot{r}^{\alpha}}
\end{aligned}
$$

After simplification, we have

$$
\begin{aligned}
& \ddot{r}^{\beta}\frac{\partial^2 L_c}{\partial \dot{r}^{\beta}\partial \dot{r}^{\alpha}} + \dot{r}^{\beta}\frac{\partial^2 L_c}{\partial r^{\beta}\partial \dot{r}^{\alpha}} - \dot{r}^{\beta}A^a_{\beta}\frac{\partial^2 L_c}{\partial s^a \partial \dot{r}^{\alpha}} - \frac{\partial L_c}{\partial r^{\alpha}} + A^b_{\alpha}\frac{\partial L_c}{\partial s^b} \\
& = \dot{r}^{\beta}\frac{\partial L}{\partial \dot{s}^b}\left(-\frac{\partial A^b_{\alpha}}{\partial r^{\beta}} + A^a_{\beta}\frac{\partial A^b_{\alpha}}{\partial s^a} + \frac{\partial A^b_{\beta}}{\partial r^{\alpha}} - A^a_{\alpha}\frac{\partial A^b_{\beta}}{\partial s^a}\right),
\end{aligned}
$$

which indeed gives the Lagrange d'Alembert equations in [BKMM]:

$$
\frac{d}{dt}\frac{\partial L_c}{\partial \dot{r}^{\alpha}} - \frac{\partial L_c}{\partial r^{\alpha}} + A^b_{\alpha}\frac{\partial L_c}{\partial s^b} = -\frac{\partial L}{\partial \dot{s}^b}B^b_{\alpha\beta}\dot{r}^{\beta}. \quad \blacksquare
$$

Remarks.

Here is another way of viewing the preceding theorem. Consider the following form of the equations:

$$
X_{\mathcal{H}} \lrcorner \, \Omega_{\mathcal{M}} = dH_{\mathcal{M}} \qquad \text{on } \mathcal{H};
$$

that is,

$$
\langle X_{\mathcal{H}} \lrcorner \, \Omega_{\mathcal{M}}, u \rangle = \langle dH_{\mathcal{M}}, u \rangle,
$$

for all $u \in \mathcal{H}$. If we rewrite this in the form

$$
\langle dH_{\mathcal{M}} - X_{\mathcal{H}} \lrcorner \, \Omega_{\mathcal{M}}, u \rangle = 0,
$$

then on the Lagrangian side, this is nothing but

$$
\langle dE_{\mathcal{D}} - X_{\mathcal{K}} \lrcorner \, (\Omega_L)_{\mathcal{D}}, v \rangle = 0,
$$

where $v \in \mathcal{K}$. With appropriate interpretations, this is equivalent to Lagrange d'Alembert principle:

$$
\begin{aligned}
\left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i}\right)(\delta q^i) &= 0 \\
\omega^a(\dot{q}) &= 0
\end{aligned}
$$

where $\omega(\delta q) = 0$.

2.5 Example: The Snakeboard

The snakeboard is a modified version of a skateboard in which the front and back pairs of wheels are independently actuated. The extra degree of freedom enables the rider to generate forward motion by twisting their body back and forth, while simultaneously moving the wheels with the proper phase relationship. For details, see [BKMM] and the references listed there. Here we will include some of the computations shown in that paper both for completeness as well as to make concrete the nonholonomic theory.

The snakeboard is modeled as a rigid body (the board) with two sets of independently actuated wheels, one on each end of the board. The human rider is modeled as a momentum wheel which sits in the middle of the board and is allowed to spin about the vertical axis. Spinning the momentum wheel causes a counter-torque to be exerted on the board. The configuration of the board is given by the position and orientation of the board in the plane, the angle of the momentum wheel, and the angles of the back and front wheels. Let (x, y, θ) represent the position and orientation of the center of the board, ψ the angle of the momentum wheel relative to the board, and ϕ_1 and ϕ_2 the angles of the back and front wheels, also relative to the board. Take the distance between the center of the board and the wheels to be r. See figure 2.5.1.

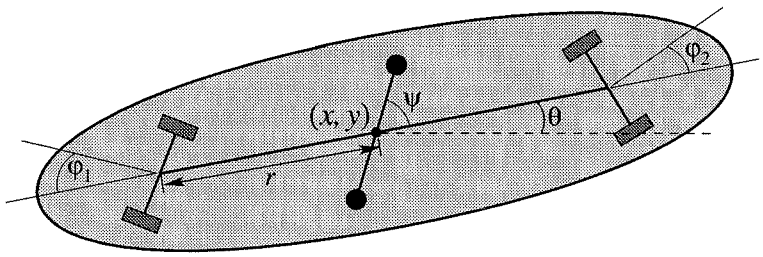

Figure 2.5.1: The geometry of the snakeboard.

In [BKMM], a simplification is made which we shall also assume in this paper, namely $\phi_1 = -\phi_2$, $J_1 = J_2$. The parameters are also chosen such that $J + J_0 + J_1 + J_2 = mr^2$, where m is the total mass of the board, J is the inertia of the board, J_0 is the inertia of the rotor and J_1, J_2 are the inertia of the wheels. This simplification eliminates some terms in the derivation but does not affect the essential geometry of the problem. Setting $\phi = \phi_1 = -\phi_2$, then the configuration space becomes $Q = SE(2) \times S^1 \times S^1$ and the Lagrangian $L : TQ \to \mathbb{R}$ is the total kinetic energy of the system and is given by

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}mr^2\dot{\theta}^2 + \frac{1}{2}J_0\dot{\psi}^2 + J_0\dot{\psi}\dot{\theta} + J_1\dot{\phi}^2.$$

The Constraints. The rolling of the front and rear wheels of the snakeboard is modeled using nonholonomic constraints which allow the wheels to spin about the vertical axis and roll in the direction that they are pointing. The wheels are not allowed to slide in the sideways direction. The constraints are defined by

$$-\sin(\theta + \phi)\dot{x} + \cos(\theta + \phi)\dot{y} - r\cos\phi\dot{\theta} = 0 \tag{2.5.1}$$

$$-\sin(\theta - \phi)\dot{x} + \cos(\theta - \phi)\dot{y} + r\cos\phi\dot{\theta} = 0 \tag{2.5.2}$$

and can be simplified as

$$\begin{aligned} \dot{x} &= -r\cot\phi\cos\theta\dot{\theta} \\ \dot{y} &= -r\cot\phi\sin\theta\dot{\theta}. \end{aligned}$$

Since the constrained Legendre transform $\mathbb{F}L|\mathcal{D}$ on the constraint submanifold $\mathcal{D}$ and its inverse are given by

$$\begin{aligned} p_x &= -mr\cot\phi\cos\theta\dot{\theta} \\ p_y &= -mr\cot\phi\sin\theta\dot{\theta} \\ p_\theta &= mr^2\dot{\theta} + J_0\dot{\psi} \\ p_\psi &= J_0\dot{\psi} + J_0\dot{\theta} \\ p_\phi &= 2J_1\dot{\phi} \\ \dot{x} &= -\frac{r}{mr^2 - J_0}\cot\phi\cos\theta(p_\theta - p_\psi) \end{aligned}$$

$$
\begin{aligned}
\dot{y} &= -\frac{r}{mr^2 - J_0} \cot\phi \sin\theta (p_\theta - p_\psi) \\
\dot{\theta} &= \frac{p_\theta - p_\psi}{mr^2 - J_0} \\
\dot{\psi} &= \frac{mr^2 p_\psi - J_0 p_\theta}{J_0(mr^2 - J_0)} \\
\dot{\phi} &= \frac{p_\phi}{2J_1},
\end{aligned}
$$

the constraint submanifold $\mathcal{M}$ is defined by

$$
\begin{aligned}
\mathcal{M} &= \{(x, y, \theta, \psi, \phi, p_x, p_y, p_\theta, p_\psi, p_\phi) \mid \\
&\quad p_x = -\frac{mr}{mr^2 - J_0} \cot\phi \cos\theta (p_\theta - p_\psi), p_y = -\frac{mr}{mr^2 - J_0} \cot\phi \sin\theta (p_\theta - p_\psi).\}
\end{aligned}
$$

Notice that $\mathcal{M}$ may be thought of as a graph in T^*Q and we can use the induced coordinates $(x, y, \theta, \psi, \phi, p_\theta, p_\psi, p_\phi)$ as its local coordinates. Hence the distribution $\mathcal{H}$ of $\mathcal{M}$ is

$$
\begin{aligned}
\mathcal{H} &= \ker\{dx + r\cot\phi\cos\theta d\theta, dy + r\cot\phi\sin\theta d\theta\} \\
&= \operatorname{span}\{-r\cot\phi\cos\theta\partial_x - r\cot\phi\sin\theta\partial_y + \partial_\theta, \partial_\psi, \partial_\phi, \partial_{p_\theta}, \partial_{p_\psi}, \partial_{p_\phi}\}.
\end{aligned}
$$

The Hamiltonian. The corresponding Hamiltonian is given via the Legendre transform by

$$
H = \frac{1}{2m}(p_x^2 + p_y^2) + \frac{1}{2J_0} p_\psi^2 + \frac{1}{2(mr^2 - J_0)}(p_\theta - p_\psi)^2 + \frac{1}{4J_1} p_\phi^2.
$$

Now if we restrict the Hamiltonian H to the submanifold $\mathcal{M}$, we get

$$
H_{\mathcal{M}} = \frac{mr^2}{2(mr^2 - J_0)^2} \cot^2\phi (p_\theta - p_\psi)^2 + \frac{1}{2J_0} p_\psi^2 + \frac{1}{2(mr^2 - J_0)}(p_\theta - p_\psi)^2 + \frac{1}{4J_1} p_\phi^2.
$$

After computing its differential $dH_{\mathcal{M}}$ and restricting it to $\mathcal{H}$, we have

$$
\begin{aligned}
dH_{\mathcal{H}} &= -\frac{mr^2}{(mr^2 - J_0)^2} \cot\phi \csc^2\phi (p_\theta - p_\psi)^2 d\phi + \frac{mr^2}{(mr^2 - J_0)^2} \cot^2\phi (p_\theta - p_\psi)(dp_\theta - dp_\psi) \\
&\quad + \frac{1}{J_0} p_\psi dp_\psi + \frac{1}{(mr^2 - J_0)}(p_\theta - p_\psi)(dp_\theta - dp_\psi) + \frac{1}{2J_1} p_\phi dp_\phi.
\end{aligned}
$$

The Two Form. After pulling back the canonical two-form of T^*Q to $\mathcal{M}$, we have

$$
\begin{aligned}
\Omega_{\mathcal{M}} &= dx \wedge dp_x + dy \wedge dp_y + d\theta \wedge dp_\theta + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi \\
&= kdx \wedge [\csc^2\phi \cos\theta (p_\theta - p_\psi) d\phi + \cot\phi \sin\theta (p_\theta - p_\psi) d\theta - \cot\phi\cos\theta(dp_\theta - dp_\psi)] \\
&\quad + kdy \wedge [\csc^2\phi \sin\theta (p_\theta - p_\psi) d\phi - \cot\phi \cos\theta (p_\theta - p_\psi) d\theta - \cot\phi\sin\theta(dp_\theta - dp_\psi)] \\
&\quad + d\theta \wedge dp_\theta + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi,
\end{aligned}
$$

where $k = mr/(mr^2 - J_0)$. If we restrict $\Omega_{\mathcal{M}}$ to the distribution $\mathcal{H}$, we get

$$
\begin{aligned}
\Omega_{\mathcal{H}} &= -kr\cot\phi\cos\theta d\theta \wedge \\
&\qquad [\csc^2\phi \cos\theta (p_\theta - p_\psi) d\phi + \cot\phi \sin\theta (p_\theta - p_\psi) d\theta - \cot\phi\cos\theta(dp_\theta - dp_\psi)] \\
&\quad - kr\cot\phi\sin\theta d\theta \wedge \\
&\qquad [\csc^2\phi \sin\theta (p_\theta - p_\psi) d\phi - \cot\phi \cos\theta (p_\theta - p_\psi) d\theta - \cot\phi\sin\theta(dp_\theta - dp_\psi)] \\
&\quad + d\theta \wedge dp_\theta + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi \\
&= d\theta \wedge [-kr\cot\phi\csc^2\phi(p_\theta - p_\psi)d\phi + kr\cot^2\phi(dp_\theta - dp_\psi) + dp_\theta] + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi.
\end{aligned}
$$

The Equations of Motion. Notice that any vector field $X_{\mathcal{M}}$ is of the form

$$X_{\mathcal{M}} = \dot{x}\partial_x + \dot{y}\partial_y + \dot{\theta}\partial_\theta + \dot{\psi}\partial_\psi + \dot{\phi}\partial_\phi + \dot{p}_\theta\partial_{p_\theta} + \dot{p}_\psi\partial_{p_\psi} + \dot{p}_\phi\partial_{p_\phi}.$$

But $X_{\mathcal{H}}$ also lies in $\mathcal{H} = \ker\{dx + r\cot\phi\cos\theta d\theta, dy + r\cot\phi\sin\theta d\theta\}$ and hence must be of the form

$$X_{\mathcal{H}} = \dot{\theta}(-r\cot\phi\cos\theta\partial_x - r\cot\phi\sin\theta\partial_y + \partial_\theta) + \dot{\psi}\partial_\psi + \dot{\phi}\partial_\phi + \dot{p}_\theta\partial_{p_\theta} + \dot{p}_\psi\partial_{p_\psi} + \dot{p}_\phi\partial_{p_\phi},$$

which gives us the first set of relationships

$$\begin{aligned} \dot{x} &= -r\cot\phi\cos\theta\,\dot{\theta} \\ \dot{y} &= -r\cot\phi\sin\theta\,\dot{\theta}. \end{aligned}$$

Moreover,

$$\begin{aligned} X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{H}} = &- kr\cot\phi\csc^2\phi(p_\theta - p_\psi)\dot{\theta}\,d\phi + kr\cot^2\phi\dot{\theta}(dp_\theta - dp_\psi) + \dot{\theta}\,dp_\theta \\ &+ \dot{\psi}\,dp_\psi + kr\cot\phi\csc^2\phi(p_\theta - p_\psi)\dot{\phi}\,d\theta + \dot{\phi}\,dp_\phi - kr\cot^2\phi\dot{p}_\theta\,d\theta - \dot{p}_\theta\,d\theta \\ &+ kr\cot^2\phi\dot{p}_\psi\,d\theta - \dot{p}_\psi\,d\psi - \dot{p}_\phi\,d\phi, \end{aligned}$$

and if equated with $dH_{\mathcal{H}}$, we will get the following set of equations:

$$kr\cot\phi\csc^2\phi(p_\theta - p_\psi)\,\dot{\phi} - kr\cot^2\phi\,\dot{p}_\theta - \dot{p}_\theta + kr\cot^2\phi\,\dot{p}_\psi = 0$$

$$\begin{aligned} -\dot{p}_\psi &= 0 \\ -\dot{p}_\phi - kr\cot\phi\csc^2\phi(p_\theta - p_\psi)\dot{\theta} &= -\frac{mr^2}{(mr^2 - J_0)^2}\cot\phi\csc^2\phi(p_\theta - p_\psi)^2 \\ kr\cot^2\phi\dot{\theta} + \dot{\theta} &= \frac{mr^2}{(mr^2 - J_0)^2}\cot^2\phi(p_\theta - p_\psi) + \frac{1}{(mr^2 - J_0)}(p_\theta - p_\psi) \\ -kr\cot^2\phi\,\dot{\theta} + \dot{\psi} &= -\frac{mr^2}{(mr^2 - J_0)^2}\cot^2\phi(p_\theta - p_\psi) + \frac{1}{J_0}p_\psi - \frac{1}{(mr^2 - J_0)}(p_\theta - p_\psi) \\ \dot{\phi} &= \frac{1}{2J_1}p_\phi. \end{aligned}$$

After simplification, we have

$$\begin{aligned} \dot{p}_\theta &= \frac{\cot\phi}{2J_1(1 - \frac{J_0}{mr^2}\sin^2\phi)}p_\phi(p_\theta - p_\psi) && (2.5.3) \\ \dot{p}_\psi &= 0 && (2.5.4) \\ \dot{p}_\phi &= 0 && (2.5.5) \\ \dot{\theta} &= \frac{p_\theta - p_\psi}{mr^2 - J_0} && (2.5.6) \\ \dot{\psi} &= \frac{mr^2 p_\psi - J_0 p_\theta}{J_0(mr^2 - J_0)} && (2.5.7) \\ \dot{\phi} &= \frac{p_\phi}{2J_1}. && (2.5.8) \end{aligned}$$

Notice that the last 3 equations are nothing but the inverse of the constrained Legendre transformation $\mathbb{F}L|\mathcal{D}$ written in local coordinates. The first equation is equivalent to the momentum equation (discussed below and in [BKMM]) written in Hamiltonian form and the 2nd and 3rd equations are the reduced equations on the shape space, again in their Hamiltonian forms.

Moreover, the corresponding Lagrangian procedure gives the equations of the motion on the Lagrangian side as

$$\ddot{\theta} - \cot\phi\dot{\phi}\dot{\theta} + \frac{J_0}{mr^2}\sin^2\phi\ddot{\psi} = 0 \quad (2.5.9)$$

$$J_0\ddot{\psi} + J_0\ddot{\theta} = 0 \quad (2.5.10)$$

$$2J_1\ddot{\phi} = 0 \quad (2.5.11)$$

and it can be shown that both systems of equations are equivalent via the Legendre transform $\mathbb{F}L|\mathcal{D}$.

3 Nonholonomic Mechanical Systems with Symmetry

Now we add the hypothesis of symmetry to the preceeding development. Assume that we have a configuration manifold Q, a Lagrangian of the form kinetic minus potential, and a distribution $\mathcal{D}$ that describes the kinematic nonholonomic constraints. We also assume there is a symmetry group G (a Lie group) that leaves the Lagrangian invariant, and that acts on Q (by isometries) and also leaves the distribution invariant, *i.e.,* the tangent of the group action maps $\mathcal{D}_q$ to $\mathcal{D}_{gq}$ (for more details, see [BKMM].) Later, we shall refer this as a **simple nonholonomic mechanical system.**

In this section,

1. We recall the basic ideas and results from [BKMM] on simple nonholonomic mechanical systems, especially on how it extend the Lagrangian reduction theory of Marsden and Scheurle [1993a,b] to the context of nonholonomic systems. We shall describe briefly how [BKMM] modifies the Ehresmann connection associated with the contraints to a new connection, called the *nonholonomic connection*, that also takes into account the symmetries, and how the reduced equations, relative to this new connection, break up into *two* sets: a set of reduced Lagrange-d'Alembert equations, and a momentum equation. When the reconstruction equations are added, one recovers the full set of equations of motion for the system.

2. We summarize the Hamiltonian reduction formulation of [BS] on finding the reduced equations of motion for nonholonomic systems with symmetry.

3. We restate the reduction procedure on the Lagrangian side corresponding to those on the Hamiltonian side using the Legendre transformation.

4. We prove that these dual procedures give us the same reduced Lagrange-d'Alembert equations in [BKMM]. Since this proof is done in coordinates, it does provide a systematic way to carry out the computations on the Hamiltonian side. Also, the proof shows where the momentum equation is lurking on the Hamiltonian side and how this is related to breaking up the dynamics of the nonholonomic system into 3 parts: a reconstruction equation for a group element g, an equation for the nonholonomic mementum p and the reduced Hamilton equations in the shape variables r, p_r (and p). This way of breaking up the dynamics may have the same significance for the control theory as what has already been noted in [BKMM].

5. We apply the Hamiltonian reduction procedure to the examples of the snakeboard, the bicycle and a nonholonomically constrained particle.

3.1 Review of Lagrangian Reduction

We first recall how [BKMM] explains in general terms how one constructs reduced systems by eliminating the group variables.

Proposition 3.1 *Under assumptions both the Lagrangian L and the distribution $\mathcal{D}$ are G-invariant, we can form the* ***reduced velocity phase space*** *TQ/G and the* ***constrained reduced velocity phase space*** *$\mathcal{D}/G$. The Lagrangian L induces well defined functions, the* ***reduced Lagrangian***

$$l : TQ/G \to \mathbb{R}$$

satisfying $L = l \circ \pi_{TQ}$ where $\pi_{TQ} : TQ \to TQ/G$ is the projection, and the ***constrained reduced Lagrangian***

$$l_c : \mathcal{D}/G \to \mathbb{R},$$

which satisfies $L|\mathcal{D} = l_c \circ \pi_{\mathcal{D}}$ where $\pi_{\mathcal{D}} : \mathcal{D} \to \mathcal{D}/G$ is the projection. Also, the Lagrange d'Alembert equations induce well defined ***reduced Lagrange d'Alembert equations*** *on $\mathcal{D}/G$. That is, the vector field on the manifold $\mathcal{D}$ determined by the Lagrange d'Alembert equations (including the constraints) is G-invariant, and so defines a reduced vector field on the quotient manifold $\mathcal{D}/G$.*

This proposition follows from general symmetry considerations, but to compute the associated reduced equations explicitly and to reconstruct the group variables, one defines the nonholonomic map J^{nh}, and extends the *NoetherTheorem* to nonholonomic system and synthesises, out of the mechanical connection and the Ehresmann connection, a nonholonomic connection $\mathcal{A}^{\mathrm{nh}}$ which is a connection on the principal bundle $Q \to Q/G$.

The Nonholonomic Momentum Map. Let the intersection of the tangent to the group orbit and the distribution at a point $q \in Q$ be denoted

$$\mathcal{S}_q = \mathcal{D}_q \cap T_q(\mathrm{Orb}(q)).$$

Define, for each $q \in Q$, the vector subspace $\mathfrak{g}^q$ to be the set of Lie algebra elements in $\mathfrak{g}$ whose infinitesimal generators evaluated at q lie in S_q:

$$\mathfrak{g}^q = \{\xi \in \mathfrak{g} \mid \xi_Q(q) \in \mathcal{S}_q\}.$$

We let $\mathfrak{g}^{\mathcal{D}}$ denote the corresponding bundle over Q whose fiber at the point q is given by $\mathfrak{g}^q$. The nonholonomic momentum map J^{nh} is the bundle map taking TQ to the bundle $(\mathfrak{g}^{\mathcal{D}})^*$ (whose fiber over the point q is the dual of the vector space $\mathfrak{g}^q$) that is defined by

$$\langle J^{\mathrm{nh}}(v_q), \xi \rangle = \frac{\partial L}{\partial \dot{q}^i}(\xi_Q)^i, \tag{3.1.1}$$

where $\xi \in \mathfrak{g}^q$. Notice that the nonholonomic momentum map may be viewed as encoding *some* of the components of the ordinary momentum map, namely the projection along those symmetry directions that are consistent with the constraints.

[BKMM] extends the Noether Theorem to nonholonomic system by deriving the equation for the momentum map that replace the usual conservation law. It is proven that if the Lagrangian L is invariant under the group action and that ξ^q is a section of the bundle $\mathfrak{g}^{\mathcal{D}}$, then any solution $q(t)$ of the Lagrange d'Alembert equations must satisfy, in addition to the given kinematic constraints, the momentum equation:

$$\frac{d}{dt}\left(J^{\mathrm{nh}}(\xi^{q(t)})\right) = \frac{\partial L}{\partial \dot{q}^i}\left[\frac{d}{dt}(\xi^{q(t)})\right]_Q^i. \tag{3.1.2}$$

When the momentum map is paired with a section in this way, we will just refer to it as the momentum. Examples show that the nonholonomic momentum map may or may not be conserved.

The Momentum Equation in Body Representation Let a local trivialization (r, g) be chosen on the principal bundle $\pi : Q \to Q/G$. Let $\eta \in \mathfrak{g}^q$ and $\xi = g^{-1}\dot{g}$. Since L is G-invariant, we can define a new function l by writing $L(r, g, \dot{r}, \dot{g}) = l(r, \dot{r}, \xi)$. Define $J^{\mathrm{nh}}_{\mathrm{loc}} : TQ/G \to (\mathfrak{g}^{\mathcal{D}})^*$ by

$$\langle J^{\mathrm{nh}}_{\mathrm{loc}}(r, \dot{r}, \xi), \eta \rangle = \left\langle \frac{\partial l}{\partial \xi}, \eta \right\rangle.$$

As with connections, J^{nh} and its version in a local trivialization are related by the Ad map; *i.e.,*

$$J^{\mathrm{nh}}(r, g, \dot{r}, \dot{g}) = \mathrm{Ad}^*_{g^{-1}} J^{\mathrm{nh}}_{\mathrm{loc}}(r, \dot{r}, \xi).$$

Choose a q-dependent basis $e_a(q)$ for the Lie algebra such that the first m elements span the subspace $\mathfrak{g}^q$. In a local trivialization, one chooses, for each r, such a basis at the identity element, say

$$e_1(r), e_2(r), \ldots, e_m(r), e_{m+1}(r), \ldots, e_k(r).$$

Define the **body fixed basis** by

$$e_a(r, g) = \mathrm{Ad}_g \cdot e_a(r);$$

thus, by G invariance, the first m elements span the subspace $\mathfrak{g}^q$. In this basis, we have

$$\langle J^{\mathrm{nh}}(r, g, \dot{r}, \dot{g}), e_b(r, g) \rangle = \left\langle \frac{\partial l}{\partial \xi}, e_b(r) \right\rangle := p_b, \tag{3.1.3}$$

which defines p_b, a function of r, $\dot{r}$ and ξ. Note that in this body representation, the functions p_b are *invariant* rather than equivariant, as is usually the case with the momentum map. It is shown in [BKMM] that in this body representation, the momentum equation is given by

$$\frac{d}{dt} p_i = \left\langle \frac{\partial l}{\partial \xi}, [\xi, e_i] + \frac{\partial e_i}{\partial r^\alpha} \dot{r}^\alpha \right\rangle, \tag{3.1.4}$$

where the range of i is 1 to m. Moreover, the momentum equation in this representation is independent of, that is, decouples from, the group variables g.

The Nonholonomic Connection Recall that in the case of simple holonomic mechanical system, the mechanical connection $\mathcal{A}$ is defined by $\mathcal{A}(v_q) = \mathbb{I}(q)^{-1} J(v_q)$ where J is the associated momentum map and $\mathbb{I}(q)$ is the locked inertia tensor of the system. Equivalently the mechanical connection can also be defined by the fact that its horizontal space at q is orthogonal to the group orbit at q with respect to the kinetic energy metric. For more information, see for example, Marsden [1992] and Marsden and Ratiu [1994].

As [BKMM] points out, in the *principal* case where the constraints and the orbit directions span the entire tangent space to the configuration space, that is,

$$\mathcal{D}_q + T_q(\mathrm{Orb}(q)) = T_q Q, \tag{3.1.5}$$

the definition of the momentum map can be used to augment the constraints and provide a connection on $Q \to Q/G$. Let J^{nh} be the nonholonomic momentum map and defined similarly as above a map $A^{\mathrm{sym}}_q : T_q Q \to \mathcal{S}_q$ given by

$$A^{\mathrm{sym}}(v_q) = (\mathbb{I}^{\mathrm{nh}}(q)^{-1} J^{\mathrm{nh}}(v_q))_Q$$

(these define the momentum "constraints") where $\mathbb{I}^{\mathrm{nh}} : \mathfrak{g}^{\mathcal{D}} \to (\mathfrak{g}^{\mathcal{D}})^*$ is the locked inertia tensor defined in a similar way as in holonomic systems.

Choose a complementary space to $\mathcal{S}_q$ by writing $T_q(\mathrm{Orb}(q)) = \mathcal{S}_q \oplus \mathcal{U}_q$. Let $A^{\mathrm{kin}}_q : T_q Q \to \mathcal{U}_q$ be a $\mathcal{U}_q$ valued form that projects $\mathcal{U}_q$ onto itself and maps $\mathcal{D}_q$ to zero. Then the kinematic constraints are defined by the equation

$$A^{\mathrm{kin}} = 0.$$

This kinematic constraints equation plus the momentum "constraints" equation can be used to synthesis a nonholonomic connection $\mathcal{A}^{\mathrm{nh}}$ which is a principal connection on the bundle $Q \to Q/G$ and whose horizontal space at the point $q \in Q$ is given by the orthogonal complement to the space $\mathcal{S}_q$ within the space $\mathcal{D}_q$. Moreover,

$$\mathcal{A}^{\mathrm{nh}}(v_q) = \mathbb{I}^{\mathrm{nh}}(q)^{-1} J^{\mathrm{nh}}(v_q). \tag{3.1.6}$$

In a body fixed basis, (3.1.6) can be written as

$$\mathrm{Ad}_g(g^{-1}\dot{g} + \mathcal{A}^{\mathrm{nh}}_{\mathrm{loc}}(r)\dot{r}) = \mathrm{Ad}_g(\mathbb{I}^{\mathrm{nh}}_{\mathrm{loc}}(r)^{-1}p). \tag{3.1.7}$$

Hence, the constraints can be represented in a nice way by

$$g^{-1}\dot{g} = \xi = -\mathcal{A}(r)\dot{r} + \Gamma(r)p, \tag{3.1.8}$$

where $\mathcal{A}(r)$ is the abbreviation for $\mathcal{A}^{\mathrm{nh}}_{\mathrm{loc}}(r)$ and $\Gamma(r) = \mathbb{I}^{\mathrm{nh}}_{\mathrm{loc}}(r)^{-1}$.

Moreover, with the help of the nonholonomic mechanical connection, the Lagrange d'Alembert principle may be broken up into two principles by breaking the variations δq into two parts, namely parts that are horizontal with respect to the nonholonomic connection and parts that are vertical (but still in $\mathcal{D}$), and the redued equations break up into *two* sets: a set of reduced Lagrange-d'Alembert equations (which have curvature terms appearing as 'forcing'), and a momentum equation, which have a form generalizing the components of the Euler-Poincaré equations along the symmetry directions consistent with the constraints. When one supplements these equations with the reconstruction equations, one recovers the full set of equations of motion for the system.

3.2 Hamiltonian Reduction

In working out the nonholonomic Hamiltonian reduction, [BS] also starts out with a *simple nonholonomic mechanical system.* Recall from Section 2 that the Legendre transformation $\mathbb{F}L : TQ \to T^*Q$ is used to define the constraint submanifold $\mathcal{M} \subset T^*Q$ where

$$\mathcal{M} = \mathbb{F}L(\mathcal{D}). \tag{3.2.1}$$

On this manifold, there is a distribution $\mathcal{H}$

$$\mathcal{H} = \mathcal{F} \cap T\mathcal{M}, \tag{3.2.2}$$

where

$$\mathcal{F} = (T\pi)^{-1}(\mathcal{D}), \tag{3.2.3}$$

and $\pi : T^*Q \to Q$. Also recall that $\Omega_{\mathcal{H}}$, the restriction of the canonical two-form Ω of T^*Q to the distribution $\mathcal{H}$ of the constraint submanifold $\mathcal{M}$, is nondegenerate and that the dynamics is given by a vector field $X_{\mathcal{H}}$ on $\mathcal{M}$ taking values in $\mathcal{H}$ and satisfies the equation

$$X_{\mathcal{H}} \lrcorner\, \Omega_{\mathcal{H}} = dH_{\mathcal{H}} \tag{3.2.4}$$

where $dH_{\mathcal{H}}$ is the (fiberwise) restriction of $dH_{\mathcal{M}}$ to $\mathcal{H}$.

Now let G be the symmetry group of this system and assume that the quotient space $\overline{\mathcal{M}} = \mathcal{M}/G$ of the G-orbit in $\mathcal{M}$ is a quotient manifold with projection map $\rho : \mathcal{M} \to \overline{\mathcal{M}}$. Since G is a symmetric group, all intrinsically defined vector fields and distributions push down to $\overline{\mathcal{M}}$. In particular, the vector field $X_{\mathcal{M}}$ on $\mathcal{M}$ pushes down to a vector field $\overline{X}_{\overline{\mathcal{M}}} = \rho_* X_{\mathcal{M}}$, and the distribution $\mathcal{H}$ pushes down to a distribution $\rho_*\mathcal{H}$ on $\overline{\mathcal{M}}$.

However, $\Omega_{\mathcal{H}}$ need not push down to a two-form defined on $\rho_*\mathcal{H}$, despite the fact that $\Omega_{\mathcal{H}}$ is G-invariant. This is because there may be infinitesimal symmetry $\xi_{\mathcal{M}}$ that lies in $\mathcal{H}$ such that

$$\xi_{\mathcal{M}} \lrcorner\, \Omega_{\mathcal{H}} \neq 0,$$

To eliminate this difficulty, [BS] restricts $\Omega_{\mathcal{H}}$ to a subdistribution $\mathcal{U}$ of $\mathcal{H}$ defined by

$$\mathcal{U} = \{u \in \mathcal{H} \mid \Omega_{\mathcal{H}}(u, v) = 0 \quad \text{for all} \quad v \in \mathcal{V} \cap \mathcal{H}\} = \mathcal{H} \cap (\mathcal{V} \cap \mathcal{H})^{\perp}, \tag{3.2.5}$$

where $\mathcal{V}$ is the distribution on $\mathcal{M}$ tangent to the orbits of G in $\mathcal{M}$ and is spanned by the infinitesimal symmetries and $(\mathcal{V} \cap \mathcal{H})^{\perp}$ is the $\Omega_{\mathcal{H}}$-orthogonal complement of $(\mathcal{V} \cap \mathcal{H})$. Clearly, $\mathcal{U}$ and $\mathcal{V}$ are both G-invariant, project down to $\overline{\mathcal{M}}$ and $\rho_* \mathcal{V} = 0$. Define $\overline{\mathcal{H}}$ by

$$\overline{\mathcal{H}} = \rho_* \mathcal{U}. \tag{3.2.6}$$

It is proven in [BS] that

1. The vector field $X_{\mathcal{H}}$ which satisfies the above Hamiltonian equation of motion (3.2.4)lies in the distribution $\mathcal{U}$.

2. The restriction $\Omega_{\mathcal{U}}$ of Ω to the distribution $\mathcal{U}$ pushes down to a nondegenerate 2-form $\Omega_{\overline{\mathcal{H}}} = \rho_* \Omega_{\mathcal{U}}$ on $\overline{\mathcal{H}}$, which is modelled by the symplectic space $(\mathcal{V} \cap \mathcal{H})^{\perp} / (\mathcal{V} \cap \mathcal{H}) \cap (\mathcal{V} \cap \mathcal{H})^{\perp}$.

3. Furthermore,
$$\overline{X}_{\overline{\mathcal{H}}} \lrcorner \, \Omega_{\overline{\mathcal{H}}} = dh_{\overline{\mathcal{H}}}, \tag{3.2.7}$$
where $h_{\overline{\mathcal{M}}} = \rho_* H_{\mathcal{M}}$ is the pushdown of the restriction to $\mathcal{M}$ of the Hamiltonian H and $dh_{\overline{\mathcal{H}}}$ is the restriction of $dh_{\overline{\mathcal{M}}}$ to $\overline{\mathcal{H}}$. This is because the equation $X_{\mathcal{H}} \lrcorner \, \Omega_{\mathcal{H}} = dH_{\mathcal{H}}$, restricted to $\mathcal{U} \subset \mathcal{H}$, vanishes on vectors in $\mathcal{V}$, and is G-invariant. Hence both sides push down to $\overline{\mathcal{H}}$.

Note that the original equations of motion are

$$X_{\mathcal{H}} \lrcorner \, \Omega_{\mathcal{H}} = dH_{\mathcal{H}} \tag{3.2.8}$$

where $\mathcal{H}$ is a distribution in the constraint manifold $\mathcal{M}$. After the reduction of symmetry we obtain equations of the same type

$$\overline{X}_{\overline{\mathcal{H}}} \lrcorner \, \Omega_{\overline{\mathcal{H}}} = dh_{\overline{\mathcal{H}}}, \tag{3.2.9}$$

where $\overline{\mathcal{H}}$ is a distribution in the reduced space $\overline{\mathcal{M}} = \mathcal{M}/G$.

3.3 Lagrangian Side

By using the Legendre transformation $\mathbb{F}L$, we can construct dual geometric structures on the tangent bundle TQ and formulate a similar Lagrangian reduction procedure. This allows us to better compare with the geometric constructions and analytic formulations on the manifold Q in [BKMM], and in the course of doing this, we realize that the requirement (see point (1) of last subsection) that the vector field $X_{\mathcal{H}}$ lies in the subdistribution $\mathcal{U}$ is equivalent to the extended Noether Theorem; that is, that any solution of the Lagrange d'Alembert equations must satisfy the momentum equation.

Recall from Section 2. W consider $\mathcal{D}$ as a constraint submanifold of TQ and then construct the distribution

$$\mathcal{K} = \mathcal{C} \cap T\mathcal{D}, \tag{3.3.1}$$

on TTQ, where

$$\mathcal{C} = (T\tau_Q)^{-1}(\mathcal{D}), \tag{3.3.2}$$

and $\tau_Q : TQ \to Q$. Clearly $\mathcal{D} = (\mathbb{F}L)^{-1}(\mathcal{M}), \mathcal{K} = (T\mathbb{F}L)^{-1}(\mathcal{H})$. The motion is then given by a vector field $X_{\mathcal{K}}$ on the manifold $\mathcal{D}$ which takes values in $\mathcal{K}$ and satisfies the equation

$$X_{\mathcal{K}} \lrcorner \, \Omega_{\mathcal{K}} = dE_{\mathcal{K}}, \tag{3.3.3}$$

where $dE_{\mathcal{K}}$ and $\Omega_{\mathcal{K}}$ are the restrictions of $dE_{\mathcal{D}}$ and $\Omega_{\mathcal{D}}$ respectively to the distribution $\mathcal{K}$.

Now let G be the symmetry group of this system and assume that the quotient space $\overline{\mathcal{D}} = \mathcal{D}/G$ of the G-orbit in $\mathcal{D}$ is a smooth quotient manifold with projection map $\lambda : \mathcal{D} \to \overline{\mathcal{D}}$. Since G is a symmetric group, all intrinsically defined vector fields and distributions push down to $\overline{\mathcal{D}}$. In particular, the vector field $X_{\mathcal{D}}$ on $\mathcal{D}$ pushes down to a vector field $\overline{X}_{\overline{\mathcal{D}}} = \lambda_* X_{\mathcal{D}}$, and the distribution $\mathcal{K}$ pushes down to a distribution $\lambda_* \mathcal{K}$ on $\overline{\mathcal{D}}$. Here we use the push forward symbol λ_* to mean that the vector fields are λ-related.

For the same reason as the Hamiltonian side, $\Omega_{\mathcal{K}}$ need not push down to a two-form defined on $\lambda_* \mathcal{K}$, despite the fact that $\Omega_{\mathcal{K}}$ is G-invariant. We can restrict $\Omega_{\mathcal{K}}$ to the subdistribution $\mathcal{W}$ of $\mathcal{K}$ defined by

$$\mathcal{W} = \{w \in \mathcal{K} \mid \Omega_{\mathcal{K}}(w, v) = 0 \text{ for all } v \in \mathcal{T} \cap \mathcal{K}\} = \mathcal{K} \cap (\mathcal{T} \cap \mathcal{D})^{\perp}, \tag{3.3.4}$$

where $\mathcal{T}$ is the distribution on $\mathcal{D}$ tangent to the orbits of G in $\mathcal{D}$ and is spanned by the infinitesimal symmetries. Clearly, $\mathcal{W}$ and $\mathcal{T}$ are both G-invariant, $\mathcal{W}$ projects down to $\overline{\mathcal{D}}$ and $\lambda_* \mathcal{T} = 0$. Define $\overline{\mathcal{K}}$ by

$$\overline{\mathcal{K}} = \lambda_* \mathcal{W}. \tag{3.3.5}$$

Since the above constructions are dual to those in the Hamiltonian side, we also have

1. The vector field $X_{\mathcal{K}}$ which satisfies the above equation (3.3.3) takes values in the distribution $\mathcal{W}$.

2. The restriction $\Omega_{\mathcal{W}}$ of Ω_L to the distribution $\mathcal{W}$, pushes down to a nondegenerate 2-form $\Omega_{\overline{\mathcal{K}}} = \lambda_* \Omega_{\mathcal{W}}$ on $\overline{\mathcal{K}}$, which is modelled by the symplectic space $(\mathcal{T} \cap \mathcal{K})^{\perp}/(\mathcal{T} \cap \mathcal{K}) \cap (\mathcal{T} \cap \mathcal{K})^{\perp}$.

3. The reduced equations of motion are given by

$$\overline{X}_{\overline{\mathcal{K}}} \lrcorner\, \Omega_{\overline{\mathcal{K}}} = d\overline{E}_{\overline{\mathcal{K}}}, \tag{3.3.6}$$

where $\overline{E}_{\overline{\mathcal{D}}} = \lambda_* E_{\mathcal{D}}$ is the pushdown of the restriction to $\mathcal{D}$ of the energy function E. This is because the equation $X_{\mathcal{K}} \lrcorner\, \Omega_{\mathcal{K}} = dE_{\mathcal{K}}$, restricted to $\mathcal{W} \subset \mathcal{K}$, vanishes on vectors in $\mathcal{T}$, and is G-invariant. Hence both sides push down to $\overline{\mathcal{K}}$. All these will become clearer in the subsequent computations.

3.4 The equivalence of Hamiltonian and Lagrangian Reductions

Theorem 3.2 *Consider a simple nonholonomic mechanical system with symmetry and assume that it is in the principal case. Then the reduction procedure on TQ described in the preceding section gives the same set of equations as in [BKMM].*

Proof The first difficulty is how to represent the constraint submanifold $\mathcal{D} \subset TQ$ in a way that is both intrinsic and ready for reduction. The comparison with the geometric constructions in [BKMM] and the desire to have the dynamics break up in a way that are ready for reconstruction give hints that we should use the tools like nonholonomic momentum p and the nonholomomic connection $\mathcal{A}$ in [BKMM] to describe the constraint submanifold $\mathcal{D}$

Recall that in [BKMM], the nonholonomic constraints together with the basic identity of the nonholonomic momentum map are used to synthesis a nonholonomic connection $\mathcal{A}$ and the nonholonomic constraints are then written in the form

$$g^{-1}\dot{g} = -A(r)\dot{r} + \Gamma(r)p, \tag{3.4.1}$$

where p is G-invariant. Hence, the constraint manifold is nothing but

$$\mathcal{D} = \{(g, r, \dot{g}, \dot{r}) \mid \dot{g} = g(-A(r)\dot{r} + \Gamma(r))p)\}. \tag{3.4.2}$$

It is a submanifold in TQ and we can use $(g, r, \dot{r}, p)$ as its induced local coordinates. Then, clearly, the corresponding coordinates for $\overline{\mathcal{D}} = \mathcal{D}/G$ are $(r, \dot{r}, p)$. From now on, we will use $A(r)$ to abbreviate $\mathcal{A}^{\text{nh}}_{\text{loc}}(r)$.

The next difficulty is to find the corresponding representations for the distribution $\mathcal{K}$, the subdistribution $\mathcal{T} \cap \mathcal{K}$ and its annihilator distribution $\mathcal{W}$ where

$$\mathcal{W} = \mathcal{K} \cap (\mathcal{T} \cap \mathcal{K})^{\perp}. \tag{3.4.3}$$

Recall that in [BKMM], a body fixed basis

$$e_b(g, r) = \text{Ad}_g \cdot e_b(r)$$

has been constructed such that the infinitesmal generators $(e_i(g, r))_Q$ of its first m elements at a point q span $\mathcal{S}_q = \mathcal{D}_q \cap T_q(\text{Orb}(q))$. Assume that G is a matrix group and e^d_i is the component of $e_i(r)$ with respect to a fixed basis $\{b_a\}$ of the Lie algebra $\mathfrak{g}$ where $(b_a)_Q = \partial_{g^a}$, then

$$(e_i(g, r))_Q = g^a_d e^d_i \partial_{g^a}.$$

Since $\mathcal{K} = (T\tau)^{-1}(\mathcal{D})$ where $\mathcal{D}_q$ is the direct sum of $\mathcal{S}_q$ and the horizontal space of the nonholonomic connection $\mathcal{A}$, it can be represented in the induced coordinates by

$$\mathcal{K} = \text{span}\{g^a_d e^d_i \partial_{g^a}, -g^a_b A^b_\alpha \partial_{g^a} + \partial_{r^\alpha}, \partial_{\dot{r}}, \partial_p\}. \tag{3.4.4}$$

Also, we have

$$\mathcal{T} \cap \mathcal{K} = \text{span}\{g^a_d e^d_i \partial_{g^a}\}. \tag{3.4.5}$$

To find the distribution $\mathcal{W}$, we have to compute $g^a_d e^d_i \partial_{g^a} \lrcorner\, \Omega_{\mathcal{D}}$, for all $i = 1, \ldots, m$. Since L is G-invariant, we have

$$\begin{aligned}
\Omega_{\mathcal{D}} &= dg^a \wedge d\left(\frac{\partial L}{\partial \dot{g}^a}\right) + dr^\alpha \wedge d\left(\frac{\partial L}{\partial \dot{r}^\alpha}\right) \\
&= dg^a \wedge d\left((g^{-1})^b_a \frac{\partial l}{\partial \xi^b}\right) + dr^\alpha \wedge d\left(\frac{\partial l}{\partial \dot{r}^\alpha}\right) \\
&= \frac{\partial (g^{-1})^b_a}{\partial g^c} \frac{\partial l}{\partial \xi^b} dg^a \wedge dg^c + (g^{-1})^b_a dg^a \wedge d\left(\frac{\partial l}{\partial \xi^b}\right) + dr^\alpha \wedge d\left(\frac{\partial l}{\partial \dot{r}^\alpha}\right)
\end{aligned}$$

Hence

$$\begin{aligned}
g^a_f e^f_i \partial_{g^a} \lrcorner\, \Omega_{\mathcal{D}} &= g^a_f e^f_i \frac{\partial (g^{-1})^b_a}{\partial g^c} \frac{\partial l}{\partial \xi^b} dg^c - g^c_f e^f_i \frac{\partial (g^{-1})^b_a}{\partial g^c} \frac{\partial l}{\partial \xi^b} dg^a + e^b_i d\left(\frac{\partial l}{\partial \xi^b}\right) \\
&= e^f_i \left(\left(g^c_f \frac{\partial (g^{-1})^b_c}{\partial g^a} - \frac{\partial (g^{-1})^b_a}{\partial g^c} g^c_f\right) \frac{\partial l}{\partial \xi^b} dg^a + d\left(\frac{\partial l}{\partial \xi^f}\right)\right) \\
&= e^f_i \left((g^{-1})^b_\sigma \left(-\frac{\partial g^\sigma_f}{\partial g^\tau} g^\tau_a + \frac{\partial g^\sigma_a}{\partial g^\tau} g^\tau_f\right) \frac{\partial l}{\partial \xi^b} (g^{-1})^a_e dg^e + d\left(\frac{\partial l}{\partial \xi^f}\right)\right) \\
&= e^f_i \left(-C^b_{af} \frac{\partial l}{\partial \xi^b} (g^{-1})^a_e dg^e + d\left(\frac{\partial l}{\partial \xi^f}\right)\right) \\
&= dp_i - \frac{\partial l}{\partial \xi^f} d(e^f_i) - C^b_{af} \frac{\partial l}{\partial \xi^b} e^f_i (g^{-1})^a_e dg^e.
\end{aligned}$$

Here, C^b_{af} is the structural constants for the Lie algebra $\mathfrak{g}$ and $p_i = \frac{\partial l}{\partial \xi^f} e^f_i$ as defined in (3.1.3). Therefore, the subdistribution $\mathcal{W} \subset \mathcal{K}$ is

$$\mathcal{W} = \ker\left\{dp_i - \frac{\partial l}{\partial \xi^f} d(e^f_i) - C^b_{af} \frac{\partial l}{\partial \xi^b} e^f_i (g^{-1})^a_e dg^e\right\}. \tag{3.4.6}$$

Since the constraint manifold $\mathcal{D}$ has the induced local coordinates $(g, r, \dot{r}, p)$, any vector field $X_{\mathcal{D}}$ on the manifold $\mathcal{D}$ is of the form

$$X_{\mathcal{D}} = \dot{g}^a \partial_{g^a} + \dot{r}^\alpha \partial_{r^\alpha} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha} + \dot{p}_i \partial_{p_i}.$$

If $X_{\mathcal{D}}$ lies in the distribution $\mathcal{K}$, then we have $\dot{g} = g(-A\dot{r} + \Gamma p)$. Moreover, if $X_{\mathcal{D}}$ lies in the distribution $\mathcal{W}$, then for each j, we have

$$\dot{p}_j - \frac{\partial l}{\partial \xi^d} \frac{\partial e_j^d}{\partial r^\alpha} \dot{r}^\alpha - C_{ad}^b \frac{\partial l}{\partial \xi^b} \xi^a e_j^d = 0,$$

i.e.,

$$\dot{p}_j = \left\langle \frac{\partial l}{\partial \xi}, [\xi, e_j] + \dot{e}_j \right\rangle, \tag{3.4.7}$$

which gives exactly the momentum equation (3.1.4). Therefore, any vector field $X_{\mathcal{W}}$ taking values in $\mathcal{W}$ must be of the form

$$X_{\mathcal{W}} = g_b^a \xi^b \partial_{g^a} + \dot{r}^\alpha \partial_{r^\alpha} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha} + \dot{p}_i \partial_{p_i}, \tag{3.4.8}$$

where

$$\xi = -A\dot{r} + \Gamma p \quad \dot{p}_j = \left\langle \frac{\partial l}{\partial \xi}, [\xi, e_j] + \dot{e}_j \right\rangle, \tag{3.4.9}$$

Now we are ready to do the reduction. But before that, we need to compute all the ingredients of the equation

$$X_{\mathcal{K}} \lrcorner \, \Omega_{\mathcal{K}} = dE_{\mathcal{K}}. \tag{3.4.10}$$

Notice first that since E is G-invariant, we have

$$\begin{aligned} E &= \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i - L \\ &= \frac{\partial L}{\partial \dot{g}^a} \dot{g}^a + \frac{\partial L}{\partial \dot{r}^\alpha} \dot{r}^\alpha - L \\ &= \frac{\partial l}{\partial \xi^a} \xi^a + \frac{\partial l}{\partial \dot{r}^\alpha} \dot{r}^\alpha - l \end{aligned}$$

After restricting it to the submanifold $\mathcal{D}$, we have

$$\begin{aligned} E_{\mathcal{D}} &= \frac{\partial l}{\partial \xi^a} (-A_\alpha^a \dot{r}^\alpha + \Gamma^{ai} p_i) + \left(\frac{\partial l_c}{\partial \dot{r}^\alpha} + A_\alpha^a \frac{\partial l}{\partial \xi^a} \right) \dot{r}^\alpha - l_c \\ &= \frac{\partial l}{\partial \xi^a} \Gamma^{ai} p_i + \frac{\partial l_c}{\partial \dot{r}^\alpha} \dot{r}^\alpha - l_c \end{aligned}$$

Therefore,

$$\begin{aligned} dE_{\mathcal{D}} &= \frac{\partial l}{\partial \xi^a} \left(\frac{\partial \Gamma^{ai}}{\partial r^\alpha} p_i dr^\alpha + \Gamma^{ai} dp_i \right) + \Gamma^{ai} p_i \left(\frac{\partial^2 l}{\partial r^\alpha \partial \xi^a} dr^\alpha + \frac{\partial^2 l}{\partial \dot{r}^\alpha \partial \xi^a} d\dot{r}^\alpha + \frac{\partial^2 l}{\partial p_j \partial \xi^a} dp_j \right) \\ &\quad + \dot{r}^\alpha \left(\frac{\partial^2 l_c}{\partial r^\beta \partial \dot{r}^\alpha} dr^\beta + \frac{\partial^2 l_c}{\partial \dot{r}^\beta \partial \dot{r}^\alpha} d\dot{r}^\beta + \frac{\partial^2 l_c}{\partial p_i \partial \dot{r}^\alpha} dp_i \right) - \frac{\partial l_c}{\partial r^\alpha} dr^\alpha - \frac{\partial l_c}{\partial p_i} dp_i. \end{aligned}$$

Furthermore,

$$X_{\mathcal{K}} \lrcorner \, \Omega_{\mathcal{D}} = g_f^a \xi^f \frac{\partial (g^{-1})_a^b}{\partial g^c} \frac{\partial l}{\partial \xi^b} dg^c - g_f^c \xi^f \frac{\partial (g^{-1})_a^b}{\partial g^c} \frac{\partial l}{\partial \xi^b} dg^a + g_f^a \xi^f (g^{-1})_a^b d \left(\frac{\partial l}{\partial \xi^b} \right)$$

$$
\begin{aligned}
& - \left(\frac{\partial}{\partial r^\alpha} \left(\frac{\partial l}{\partial \xi^b} \right) \dot{r}^\alpha + \frac{\partial}{\partial \dot{r}^\alpha} \left(\frac{\partial l}{\partial \xi^b} \right) \ddot{r}^\alpha + \frac{\partial}{\partial p_i} \left(\frac{\partial l}{\partial \xi^b} \right) \dot{p}_i \right) (g^{-1})^b_a dg^a \\
& + (\dot{r}^\alpha \partial_{r^\alpha} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha} + \dot{p}_i \partial p_i) \lrcorner \left(dr^\alpha \wedge d \left(\frac{\partial l}{\partial \dot{r}^\alpha} \right) \right) \\
= \; & \xi^f d \left(\frac{\partial l}{\partial \xi^f} \right) + \left(C^b_{fa} \frac{\partial l}{\partial \xi^b} \xi^f - \frac{d}{dt} \left(\frac{\partial l}{\partial \xi^a} \right) \right) (g^{-1})^a_e dg^e \\
& + (\dot{r}^\alpha \partial_{r^\alpha} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha} + \dot{p}_i \partial p_i) \lrcorner \left(dr^\alpha \wedge d \left(\frac{\partial l}{\partial \dot{r}^\alpha} \right) \right). \qquad (3.4.11)
\end{aligned}
$$

Clearly, both sides of the equation

$$X_{\mathcal{K}} \lrcorner \Omega_{\mathcal{K}} = dE_{\mathcal{K}} \qquad (3.4.12)$$

are G-invariant, and when restricted to subdistribution $\mathcal{W} \subset \mathcal{K}$, they vanish on the distribution $\mathcal{T} \cap \mathcal{K}$. This can be shown to be true either by invoking how $\mathcal{W}$ has been constructed or by direct calculation, noticing that when

$$\left(C^b_{fa} \frac{\partial l}{\partial \xi^b} \xi^f - \frac{d}{dt} \left(\frac{\partial l}{\partial \xi^a} \right) \right) (g^{-1})^a_e dg^e \qquad (3.4.13)$$

is paired with $g^f_c e^c_i$ in $\mathcal{T} \cap \mathcal{K}$, it is equal to zero on $\mathcal{W}$. Hence both sides push down to $\overline{\mathcal{K}}$ where

$$\overline{X_{\overline{\mathcal{K}}}} = \dot{r}^\alpha \partial_{r^\alpha} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha} + \dot{p}_i \partial_{p_i}, \qquad (3.4.14)$$

with

$$\dot{p}_i = \left\langle \frac{\partial l}{\partial \xi}, [\xi, e_i] + \dot{e}_i \right\rangle. \qquad (3.4.15)$$

To find the remaining reduced equations, notice that the restriction of (3.4.13) to the subdistribution spanned by $\{-g^a_b A^b_\alpha \partial_{g^a} + \partial_{r^\alpha}, \partial_{\dot{r}^\alpha}, \partial_{p_i}\}$ is equivalent to

$$- \left(C^b_{fa} \frac{\partial l}{\partial \xi^b} \xi^f - \frac{d}{dt} \left(\frac{\partial l}{\partial \xi^a} \right) \right) A^a_\alpha dr^\alpha. \qquad (3.4.16)$$

If we compute

$$- \left(C^b_{fa} \frac{\partial l}{\partial \xi^b} \xi^f - \frac{d}{dt} \left(\frac{\partial l}{\partial \xi^a} \right) \right) A^a_\alpha dr^\alpha + \xi^a d \left(\frac{\partial l}{\partial \xi^a} \right) + (\dot{r}^\alpha \partial_{r^\alpha} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha} + \dot{p}_i \partial p_i) \lrcorner \left(dr^\alpha \wedge d \left(\frac{\partial l}{\partial \dot{r}^\alpha} \right) \right)$$

and equate its terms with the corresponding terms of $d\overline{E}_{\overline{\mathcal{K}}}$ which is the same as $dE_{\mathcal{K}}$, we have the following equations after some computations

$$\frac{d}{dt} \left(\frac{\partial l_c}{\partial \dot{r}^\alpha} \right) - \frac{\partial l_c}{\partial r^\alpha} = -C^b_{da} \frac{\partial l}{\partial \xi^b} \xi^d A^a_\alpha - \frac{\partial l}{\partial \xi^a} \left(\dot{A}^a_\alpha - \frac{\partial A^a_\beta}{\partial r^\alpha} \dot{r}^\beta + \frac{\partial \Gamma^{ai} p_i}{\partial r^\alpha} \right).$$

After plugging in the constraint $\xi = -A\dot{r} + \Gamma p$ and simplify, we get the desired reduced equations

$$\frac{d}{dt} \left(\frac{\partial l_c}{\partial \dot{r}^\alpha} \right) - \frac{\partial l_c}{\partial r^\alpha} = -\frac{\partial l}{\partial \xi^b} (B^b_{\alpha\beta} \dot{r}^\beta + F^{bi} p_i), \qquad (3.4.17)$$

where

$$
\begin{aligned}
B^b_{\alpha\beta} &= \frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} - C^b_{ac} A^a_\beta A^c_\alpha \qquad (3.4.18) \\
F^{bi}_\alpha &= \frac{\partial \Gamma^{bi}}{\partial r^\alpha} - C^b_{ad} A^a_\alpha \Gamma^{di}. \qquad (3.4.19)
\end{aligned}
$$

In an orthogonal body frame where we choose our moving basis $e_b(g,r)$ to be orthogonal, that is, the corresponding generators $[e_b(g,r)]_Q$ are orthogonal in the given kinetic energy metric (actually, all that is needed is that the vectors in the set of basis vectors corresponding to the subspace $\mathcal{S}_q$ be orthogonal to the remaining basis vectors), the momentum equation (3.4.7) can be written as (see BKMM)

$$\frac{d}{dt}p_i = C^j_{hi} I^{hl} p_j p_l + \mathcal{D}^j_{i\alpha}\dot{r}^\alpha p_j + \mathcal{D}_{\alpha\beta i}\dot{r}^\alpha\dot{r}^\beta, \tag{3.4.20}$$

where

$$\begin{aligned}
\mathcal{D}^j_{i\alpha} &= -C^j_{ai}A^a_\alpha + \gamma^j_{i\alpha} + \lambda_{a'\alpha}C^{a'}_{li}I^{lj} && (3.4.21)\\
\mathcal{D}_{\alpha\beta i} &= \lambda_{a'\alpha}(-C^{a'}_{ai}A^a_\beta + \gamma^{a'}_{i\beta}). && (3.4.22)
\end{aligned}$$

Here $\gamma^c_{b\alpha}$ and $\lambda_{a'\alpha}$ are defined by

$$\begin{aligned}
\frac{\partial e_b}{\partial r^\alpha} &= \gamma^c_{b\alpha}e_c && (3.4.23)\\
\lambda_{a'\alpha} &= \frac{\partial l}{\partial\xi^{a'}\partial\dot{r}^\alpha} - \frac{\partial l}{\partial\xi^{a'}\partial\xi^b}A^b_\alpha. && (3.4.24)
\end{aligned}$$

Notice that while the summation range of $a, b, c, d...$ are over all Lie algebra element (1 to k). those over $i, j, l, ...$ are the restricted (constrained) range (1 to m) and those over $a', b', ...$ run from $m+1$ to k (which correspond to the symmetry directions not aligned with the constraints).

Similarly we can rewrite the above reduced Hamilton equations (3.4.17) using the orthogonal body frame. Essentially, it is a change of basis. Instead of using the natural fixed basis $\{b_a\}$ where $(b_a)_Q = \partial_{g^a}$, we do all the computations in the orthogonal body frame $\{e_b\}$. With the abuse of notation, we shall still use $l(r,\dot{r},\xi)$ and $l_c(r,\dot{r},p)$ to denote the reduced Lagrangian and the constrained reduced Lagrangian (in the orthogonal body frame) respectively. But it should be clear that for the following computations, $\xi = \xi^b e_b$. Similar interpretation should apply to all other notations. Now let us compute the right hand side of the reduced equations in the new basis. Since

$$\begin{aligned}
\frac{\partial}{\partial r^\beta}(A^a_\alpha e_a) &= \frac{\partial A^b_\alpha}{\partial r^\beta}e_b + A^a_\alpha\gamma^b_{a\beta}e_b\\
\frac{\partial}{\partial r^\alpha}(\Gamma^{bi}p_i e_b) &= \frac{\partial\Gamma^{bi}}{\partial r^\alpha}p_i e_b + \Gamma^{ai}\gamma^b_{a\alpha}p_i e_b,
\end{aligned}$$

we have

$$\begin{aligned}
\frac{d}{dt}\left(\frac{\partial l_c}{\partial\dot{r}^\alpha}\right) - \frac{\partial l_c}{\partial r^\alpha} &= -\frac{\partial l}{\partial\xi^b}\left(\frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} - C^b_{ac}A^a_\beta A^c_\alpha + A^c_\alpha\gamma^b_{c\beta} - A^c_\beta\gamma^b_{c\alpha}\right)\dot{r}^\beta\\
&\quad -\frac{\partial l}{\partial\xi^b}\left(\frac{\partial\Gamma^{bi}p_i}{\partial r^\alpha} - C^b_{ad}A^a_\alpha\Gamma^{di}p_i + \Gamma^{ci}\gamma^b_{c\alpha}p_i\right)
\end{aligned}$$

Now applying Proposition 7.1 of [BKMM] to the above reduced equations and notice that in the orthogonal body basis, $\Gamma^{bi} = 0$ for any $b > m$ (recall $\Gamma^{ji} = I^{ji}$), we can re-write the reduced equations in the following form

$$\frac{d}{dt}\left(\frac{\partial l_c}{\partial\dot{r}^\alpha}\right) - \frac{\partial l_c}{\partial r^\alpha} = -(\mathcal{K}^{jl}_\alpha p_j p_l + \mathcal{K}^j_{\alpha\beta}\dot{r}^\beta p_j + \mathcal{K}_{\alpha\beta\delta}\dot{r}^\beta\dot{r}^\delta) \tag{3.4.25}$$

where

$$\begin{aligned}
\mathcal{K}^{jl}_\alpha &= \frac{\partial I^{jl}}{\partial r^\alpha} - C^j_{bh}A^b_\alpha I^{hl} + \gamma^j_{h\alpha}I^{hl} && (3.4.26)\\
\mathcal{K}^j_{\alpha\beta} &= \lambda_{a'\beta}(-C^{a'}_{bh}A^b_\alpha I^{hj} + \gamma^{a'}_{h\alpha}I^{hj}) + B^j_{\alpha\beta} && (3.4.27)\\
\mathcal{K}_{\alpha\beta\delta} &= \lambda_{a'\delta}B^{a'}_{\alpha\beta}. && (3.4.28)
\end{aligned}$$

Here

$$B_{\alpha\beta}^b = \frac{\partial A_\alpha^b}{\partial r^\beta} - \frac{\partial A_\beta^b}{\partial r^\alpha} - C_{ac}^b A_\beta^a A_\alpha^c + A_\alpha^c \gamma_{c\beta}^b - A_\beta^c \gamma_{c\alpha}^b. \quad \blacksquare \tag{3.4.29}$$

Remarks

1. A careful reading of the proof of Theorem 3.2 and the subsections 3.2 and 3.3 shows that the Hamiltonian reduction procedure still works as long as the constrained Legendre transform $\mathbb{F}L|\mathcal{D}$ is invertible. This is important because in some examples like the bicycle the Legendre transform $\mathbb{F}L$ is singular, but its restriction to the constraint submanifold $\mathcal{D}$ is invertible and the Hamiltonian reduction prodecure is also applicable.

2. In many examples like the snakeboard and the bicycle, the constraints satisfy a special condition, namely, they involve only the velocities of the group variables $\dot{g}$ and are independent of the velocities of the shape variables $\dot{r}$ (see equations (2.5.1) and (2.5.2)). Under this special condition, the distribution $\mathcal{K}$ in equation (3.4.4) can be represented by

$$\mathcal{K} = \operatorname{span}\{g_d^a e_i^d \partial_{g^a}, \partial_r, \partial_{\dot{r}}, \partial_p\}. \tag{3.4.30}$$

This representation simplifies the computation for finding the reduced equations because the restriction of the one form (3.4.13) to the subdistribution $\overline{\mathcal{K}}$ spanned by $\{\partial_r, \partial_{\dot{r}}, \partial_p\}$ will equal to zero. Hence in pushing down $X_{\mathcal{K}} \lrcorner\, \Omega_{\mathcal{D}}$ in (3.4.11) to $\overline{\mathcal{K}}$, we can simply omit the one form (3.4.13). In the following subsections, we will use this simplified procedure for the examples of the snakeboard and the bicycle. We will use a modified version of a nonholonomically constrained particle to illustrate the general procedure.

3. Since the momentum equation is central to the theory of nonholonomic mechanical systems with symmetry, we make a few additional remarks about it. Before that, we state the following proposition, the result of which is implicit in both [BKMM] and Ostrowski [1996].

Proposition 3.3 *For a nonholonomic mechanical system with symmetry, we have*

$$\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i}\right)(\xi_Q^q)^i = \frac{d}{dt}\left(\left(\frac{\partial L}{\partial \dot{q}^i}\right)(\xi_Q^q)^i\right) - \frac{\partial L}{\partial \dot{q}^i}\left(\frac{d}{dt}\xi^q\right)_Q^i \tag{3.4.31}$$

where $\xi^q \in \mathfrak{g}^q$

Proof: Choose a section of $\mathfrak{g}^{\mathcal{D}}$ and apply the chain rule to give

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}(\xi_Q^q)^i\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right)(\xi_Q^q)^i + \frac{\partial L}{\partial \dot{q}^i}\left((T\xi_Q^q \cdot \dot{q})^i + \left(\frac{d}{dt}\xi^q\right)_Q^i\right)$$

Invariance of the Lagrangian implies that

$$L(exp(s\xi^q) \cdot q, \exp(s\xi^q) \cdot \dot{q}) = L(q, \dot{q}).$$

Differentiating this expression and evaluating it at $s = 0$, we get

$$\frac{\partial L}{\partial q^i}(\xi_Q^q)^i + \frac{\partial L}{\partial \dot{q}^i}(T\xi_Q^q \cdot \dot{q})^i = 0$$

After eliminating the term $\frac{\partial L}{\partial \dot{q}^i}(T\xi_Q^q \cdot \dot{q})^i$ from the above two equations, we arrive at the desired result. $\blacksquare$

The above equation can be rewritten as

$$\left\langle (dE - X \lrcorner \Omega_L)|_{\mathcal{D}}, (\xi_Q^q)' \right\rangle = \frac{d}{dt}\left(\left(\frac{\partial L}{\partial \dot{q}^i}\right)(\xi_Q^q)^i\right) - \frac{\partial L}{\partial \dot{q}^i}\left(\frac{d}{dt}\xi^q\right)_Q^i, \tag{3.4.32}$$

where $(\xi_Q^q)' \in \mathcal{T} \cap \mathcal{K}$ and $T\tau_Q((\xi_Q^q)') = \xi_Q^q$. Since both the energy function E and the submanifold $\mathcal{D}$ are G-invariant, the left hand of the above equation reduces to $\Omega_{\mathcal{D}}(X_{\mathcal{D}}, (\xi_Q^q)')$ and hence any vector field $X_{\mathcal{D}}$ which takes values in $\mathcal{W} = \mathcal{K} \cap (\mathcal{T} \cap \mathcal{K})^{\perp}$ will make the left hand side zero and hence must satisfy the momentum equation (3.1.2)

$$\frac{d}{dt}\left(\left(\frac{\partial L}{\partial \dot{q}^i}\right)(\xi_Q^q)^i\right) - \frac{\partial L}{\partial \dot{q}^i}\left(\frac{d}{dt}\xi^q\right)_Q^i = 0, \tag{3.4.33}$$

as we have already seen in the proof of Theorem 3.2.

In showing that the vector field $X_{\mathcal{H}}$ which satisfies the equation $X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{H}} = dH_{\mathcal{H}}$ must lie in the subdistribution $\mathcal{U}$, one might think that any vector field $Y \in \mathcal{V} \cap \mathcal{H}$ can be expressed as a linear combination of infinitesmal generators (generated by fixed Lie algebra elements). But this is not the case, as we have pointed out earlier in the Lagrangian side, in general $(\xi_Q^q)'$ is the (vertical) lift of a section of the bundle $\mathcal{S}$ (generated by a section of the bundle $\mathfrak{g}^{\mathcal{D}}$). This is also true on the Hamiltonian side.

3.5 Example: The Snakeboard Revisited

Now we return to the snakeboard and discuss the role of the symmetry group $G = SE(2)$. Recall from our earlier discussion that the Lagrangian is

$$L(q, \dot{q}) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}mr^2\dot{\theta}^2 + +\frac{1}{2}J_0\dot{\psi}^2 + J_0\dot{\psi}\dot{\theta} + J_1\dot{\phi}_1^2, \tag{3.5.1}$$

which is independent of the configuration of the board and hence it is invariant to all possible group actions.

The Constraint Submanifold. The condition of rolling without slipping gives rise to the constraint one forms

$$\begin{aligned} \omega_1(q) &= -\sin(\theta + \phi)dx + \cos(\theta + \phi)dy - r\cos\phi d\theta \\ \omega_2(q) &= -\sin(\theta - \phi)dx + \cos(\theta - \phi)dy + r\cos\phi d\theta, \end{aligned}$$

which are invariant under the $SE(2)$ action. The constraints determine the kinematic distribution $\mathcal{D}_q$:

$$\mathcal{D}_q = \operatorname{span}\{\partial_\psi, \partial_\phi, a\partial_x + b\partial_y + c\partial_\theta\},$$

where $a = -2r\cos^2\phi\cos\theta, b = -2r\cos^2\phi\sin\theta, c = \sin 2\phi$. The tangent space to the orbits of the $SE(2)$ action is given by

$$T_q(\mathrm{Orb}(q)) = \operatorname{span}\{\partial_x, \partial_y, \partial_\theta\}$$

The intersection between the tangent space to the group orbits and the constraint distribution is thus given by

$$\mathcal{S}_q = \mathcal{D}_q \cap T_q(\mathrm{Orb}(q)) = \operatorname{span}\{a\partial_x + b\partial_y + c\partial_\theta\}.$$

The momentum can be constructed by choosing a section of $\mathcal{S} = \mathcal{D} \cap T\mathrm{Orb}$ regarded as a bundle over Q. Since $\mathcal{D}_q \cap T_q\mathrm{Orb}(q)$ is one-dimensional, the section can be chosen to be

$$\xi_Q^q = a\partial_x + b\partial_y + c\partial_\theta,$$

which is invariant under the action of $SE(2)$ on Q. The nonholonomic momentum is thus given by

$$\begin{aligned} p &= \frac{\partial L}{\partial \dot{q}^i}(\xi_Q^q)^i \\ &= ma\dot{x} + mb\dot{y} + mr^2 c\dot{\theta} + J_0 c\dot{\psi}. \end{aligned}$$

The kinematic constraints plus the momentum are given by

$$\begin{aligned} 0 &= -\sin(\theta+\phi)\dot{x} + \cos(\theta+\phi)\dot{y} - r\cos\phi\dot{\theta} \\ 0 &= -\sin(\theta-\phi)\dot{x} + \cos(\theta-\phi)\dot{y} + r\cos\phi\dot{\theta} \\ p &= -2mr\cos^2\phi\cos\theta\dot{x} - 2mr\cos^2\phi\sin\theta\dot{y} \\ &\quad + mr^2\sin 2\phi\dot{\theta} + J_0\sin 2\phi\dot{\psi}. \end{aligned}$$

Adding, subtracting, and scaling these equations, we can write (away from the point $\phi = \pi/2$),

$$\begin{bmatrix} \cos\theta\dot{x} + \sin\theta\dot{y} \\ -\sin\theta\dot{x} + \cos\theta\dot{y} \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} -\dfrac{J_0}{2mr}\sin 2\phi\dot{\psi} \\ 0 \\ \dfrac{J_0}{mr^2}\sin^2\phi\dot{\psi} \end{bmatrix} = \begin{bmatrix} \dfrac{-1}{2mr}p \\ 0 \\ \dfrac{\tan\phi}{2mr^2}p \end{bmatrix} \tag{3.5.2}$$

These equations have the form

$$g^{-1}\dot{g} + A(r)\dot{r} = \Gamma(r)p$$

where

$$\begin{aligned} A(r) &= -\frac{J_0}{2mr}\sin 2\phi e_x\, d\psi + \frac{J_0}{mr^2}\sin^2\phi e_\theta\, d\psi \\ \Gamma(r) &= \frac{-1}{2mr}e_x + \frac{1}{2mr^2}\tan\phi\, e_\theta. \end{aligned}$$

These are precisely the terms which appear in the nonholonomic connection relative to the (global) trivialization (r, g).

After applying the constrained Legendre transformation and its inverse to the constraint equations (3.5.2), we have

$$\begin{bmatrix} \cos\theta p_x + \sin\theta p_y \\ -\sin\theta p_x + \cos\theta p_y \\ p_\theta \end{bmatrix} + \begin{bmatrix} -\dfrac{mr\sin\phi\cos\phi}{(mr^2 - J_0\sin^2\phi)}p_\psi \\ 0 \\ -\dfrac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)}p_\psi \end{bmatrix} = \begin{bmatrix} \dfrac{-mr}{2(mr^2 - J_0\sin^2\phi)}p \\ 0 \\ \dfrac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)}p \end{bmatrix}, \tag{3.5.3}$$

where

$$p = -2r\cos^2\phi\cos\theta p_x - 2r\cos^2\phi\sin\theta p_y + \sin 2\phi p_\theta$$

and is $SE(2)$-invariant.

Therefore, the constraint submanifold $\mathcal{M} \subset T^*Q$ is defined by

$$\begin{aligned} p_x &= \frac{mr\sin\phi\cos\phi}{(mr^2 - J_0\sin^2\phi)}p_\psi\cos\theta - \frac{mr}{2(mr^2 - J_0\sin^2\phi)}p\cos\theta \\ p_y &= \frac{mr\sin\phi\cos\phi}{(mr^2 - J_0\sin^2\phi)}p_\psi\sin\theta - \frac{mr}{2(mr^2 - J_0\sin^2\phi)}p\sin\theta \\ p_\theta &= \frac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)}p_\psi + \frac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)}p \end{aligned}$$

It is a submanifold in T^*Q and we can use $(x, y, \theta, \psi, \phi, p_\psi, p_\phi, p)$ as its induced local coordinates.

The Distributions $\mathcal{H}, \mathcal{V}\cap\mathcal{H}$ and $\mathcal{U}$. With the induced coordinates, the distribution $\mathcal{H}$ on $\mathcal{M}$ is

$$\mathcal{H} = \mathrm{span}\{-2r\cos^2\phi\cos\theta\partial_x - 2r\cos^2\phi\sin\theta\partial_y + \sin 2\phi\partial_\theta, \partial_\psi, \partial_\phi, \partial_{p_\psi}, \partial_{p_\phi}, \partial_p\} \tag{3.5.4}$$

and the subdistribution $\mathcal{V}\cap\mathcal{H}$ is

$$\mathcal{V}\cap\mathcal{H} = \mathrm{span}\{-2r\cos^2\phi\cos\theta\partial_x - 2r\cos^2\phi\sin\theta\partial_y + \sin 2\phi\partial_\theta\}. \tag{3.5.5}$$

As for the subdistribution $\mathcal{U}$, we first calculate the two form $\Omega_{\mathcal{M}}$. After pulling back the canonical two-form of T^*Q to $\mathcal{M}$, we have

$$\begin{aligned}
\Omega_{\mathcal{M}} &= dx\wedge dp_x + dy\wedge dp_y + d\theta\wedge dp_\theta + d\psi\wedge dp_\psi + d\phi\wedge dp_\phi \\
&= (\cos\theta dx + \sin\theta dy)\wedge\left(\frac{mr\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}dp_\psi - \frac{mr}{2(mr^2 - J_0\sin^2\phi)}dp\right) \\
&\quad + (\cos\theta dx + \sin\theta dy)\wedge\left(\frac{mr(mr^2\cos 2\phi + J_0\sin^2\phi)}{(mr^2 - J_0\sin^2\phi)^2}p_\psi d\phi - \frac{mrJ_0\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)^2}pd\phi\right) \\
&\quad + d\theta\wedge\left(\frac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)}dp_\psi + \frac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)}dp\right) \\
&\quad + d\theta\wedge\left(\frac{mr^2(J_0 - mr^2)\sin 2\phi}{(mr^2 - J_0\sin^2\phi)^2}p_\psi d\phi + \frac{(mr^2 - J_0)(mr^2\sec^2\phi + J_0\tan^2\phi\cos 2\phi)}{2(mr^2 - J_0\sin^2\phi)^2}pd\phi\right) \\
&\quad + (-\sin\theta dx + \cos\theta dy)\wedge\left(\frac{mr\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}p_\psi - \frac{mr}{2(mr^2 - J_0\sin^2\phi)}p\right)d\theta \\
&\quad + d\psi\wedge dp_\psi + d\phi\wedge dp_\phi
\end{aligned}$$

Since $\mathcal{U} = (\mathcal{V}\cap\mathcal{H})^\perp = \ker\{(\mathcal{V}\cap\mathcal{H}) \lrcorner\, \Omega_{\mathcal{H}}\}$, we need to calculate $(\mathcal{V}\cap\mathcal{H}) \lrcorner\, \Omega_{\mathcal{M}}$, and restrict it to $\mathcal{H}$:

$$\begin{aligned}
&(\mathcal{V}\cap\mathcal{H}) \lrcorner\, \Omega_{\mathcal{H}} = \\
&\quad -2r\cos^2\phi\left(\frac{mr\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}dp_\psi - \frac{mr}{2(mr^2 - J_0\sin^2\phi)}dp\right) \\
&\quad -2r\cos^2\phi\left(\frac{mr(mr^2\cos 2\phi + J_0\sin^2\phi)}{(mr^2 - J_0\sin^2\phi)^2}p_\psi d\phi - \frac{mrJ_0\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)^2}pd\phi\right) \\
&\quad + \sin 2\phi\left(\frac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)}dp_\psi + \frac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)}dp\right) \\
&\quad + \sin 2\phi\left(\frac{mr^2(J_0 - mr^2)\sin 2\phi}{(mr^2 - J_0\sin^2\phi)^2}p_\psi d\phi + \frac{(mr^2 - J_0)(mr^2\sec^2\phi + J_0\tan^2\phi\cos 2\phi)}{2(mr^2 - J_0\sin^2\phi)^2}pd\phi\right) \\
&= dp - \frac{2mr^2\cos^2\phi}{mr^2 - J_0\sin^2\phi}p_\psi d\phi + \frac{(mr^2 + J_0\cos 2\phi)\tan\phi}{mr^2 - J_0\sin^2\phi}pd\phi
\end{aligned}$$

Hence,

$$\mathcal{U} = \ker\left\{dp - \frac{2mr^2\cos^2\phi}{mr^2 - J_0\sin^2\phi}p_\psi d\phi + \frac{(mr^2 + J_0\cos 2\phi)\tan\phi}{mr^2 - J_0\sin^2\phi}pd\phi\right\}. \tag{3.5.6}$$

The Reconstruction and Momentum Equations A vector field $X_{\mathcal{U}}$ taking values in $\mathcal{U}$ must be of the form

$$X_{\mathcal{U}} = \dot{x}\partial_x + \dot{y}\partial_y + \dot{\theta}\partial_\theta + \dot{\psi}\partial_\psi + \dot{\phi}\partial_\phi + \dot{p}_\psi\partial_{p_\psi} + \dot{p}_\phi\partial_{p_\phi} + \dot{p}\partial_p \tag{3.5.7}$$

where

$$\begin{aligned}
\dot{x} &= \frac{J_0}{2mr}\sin 2\phi\dot{\psi}\cos\theta - \frac{1}{2mr}p\cos\theta \\
\dot{y} &= \frac{J_0}{2mr}\sin 2\phi\dot{\psi}\sin\theta - \frac{1}{2mr}p\sin\theta \\
\dot{\theta} &= -\frac{J_0}{mr^2}\sin^2\phi\dot{\psi} + \frac{\tan\phi}{2mr^2}p
\end{aligned}$$

and

$$\dot{p} = \frac{2mr^2\cos^2\phi}{mr^2 - J_0\sin^2\phi}p_\psi\dot{\phi} - \frac{(mr^2 + J_0\cos 2\phi)\tan\phi}{mr^2 - J_0\sin^2\phi}p\dot{\phi} \tag{3.5.8}$$

The equations for $\dot{x}$, $\dot{y}$ and $\dot{\theta}$ are the same reconstruction equations as equations (3.5.2) and the last one for $\dot{p}$ is the momentum eqution on the Hamiltonian side. As noted in [BKMM], the momentum p is the angular momentum of the system about the point P shown in figure 3.5.1.

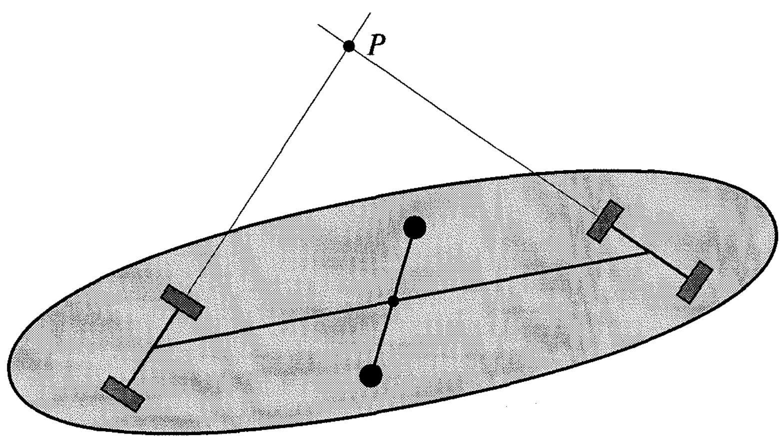

Figure 3.5.1: The momentum p is the angular momentum of the snakeboard system about the point P.

It can be checked that the momentum equation (3.5.8) is equivalent to the equation (2.5.3) via a change of variables with

$$\begin{aligned}
p &= -2r\cos^2\phi\cos\theta p_x - 2r\cos^2\phi\sin\theta p_y + \sin 2\phi p_\theta \\
&= \frac{2(mr^2 - J_0\sin^2\phi)\cot\phi}{mr^2 - J_0}p_\theta - \frac{2mr^2\cos^2\phi\cot\phi}{mr^2 - J_0}p_\psi
\end{aligned}$$

as the key link. Similarly the two full sets of equations of motion in both section 2.5 and this section are also related in the same way.

The Reduced Hamilton Equations. To find the remaining reduced equations, we need to compute

$$X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}} = dH_{\mathcal{M}}, \tag{3.5.9}$$

restrict it to the subdistribution $\mathcal{U}$ and then push it down to the reduced constraint submanifold $\overline{\mathcal{M}}$. Let us first compute $X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}}$

$X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}} =$

$$
\begin{aligned}
&(\dot{x}\cos\theta+\dot{y}\sin\theta)\left(\frac{mr\sin 2\phi}{2(mr^2-J_0\sin^2\phi)}dp_\psi-\frac{mr}{2(mr^2-J_0\sin^2\phi)}dp\right)\\
&+(\dot{x}\cos\theta+\dot{y}\sin\theta)\left(\frac{mr(mr^2\cos 2\phi+J_0\sin^2\phi)}{(mr^2-J_0\sin^2\phi)^2}p_\psi d\phi-\frac{mrJ_0\sin 2\phi}{2(mr^2-J_0\sin^2\phi)^2}pd\phi\right)\\
&+\dot{\theta}\left(\frac{mr^2\cos^2\phi}{(mr^2-J_0\sin^2\phi)}dp_\psi+\frac{(mr^2-J_0)\tan\phi}{2(mr^2-J_0\sin^2\phi)}dp\right)\\
&+\dot{\theta}\left(\frac{mr^2(J_0-mr^2)\sin 2\phi}{(mr^2-J_0\sin^2\phi)^2}p_\psi d\phi+\frac{(mr^2-J_0)(mr^2\sec^2\phi+J_0\tan^2\phi\cos 2\phi)}{2(mr^2-J_0\sin^2\phi)^2}pd\phi\right)\\
&+\dot{\psi}dp_\psi+\dot{\phi}dp_\phi-\dot{p}_\psi d\psi-\dot{p}_\phi d\phi\\
&-\dot{\theta}\left(\frac{mr\sin 2\phi}{2(mr^2-J_0\sin^2\phi)}p_\psi-\frac{mr}{2(mr^2-J_0\sin^2\phi)}p\right)(-\sin\theta dx+\cos\theta dy)\\
&-mr\left(\frac{mr^2\cos 2\phi+J_0\sin^2\phi}{(mr^2-J_0\sin^2\phi)^2}p_\psi\dot{\phi}-\frac{J_0\sin 2\phi}{2(mr^2-J_0\sin^2\phi)^2}p\dot{\phi}\right)(\cos\theta dx+\sin\theta dy)\\
&-\left(\frac{mr^2(J_0-mr^2)\sin 2\phi}{(mr^2-J_0\sin^2\phi)^2}p_\psi\dot{\phi}+\frac{(mr^2-J_0)(mr^2\sec^2\phi+J_0\tan^2\phi\cos 2\phi)}{2(mr^2-J_0\sin^2\phi)^2}p\dot{\phi}\right)d\theta\\
&-\frac{mr\sin 2\phi}{2(mr^2-J_0\sin^2\phi)}\dot{p}_\psi(\cos\theta dx+\sin\theta dy)-\frac{mr^2\cos^2\phi}{(mr^2-J_0\sin^2\phi)}\dot{p}_\psi d\theta\\
&+\frac{mr}{2(mr^2-J_0\sin^2\phi)}(\cos\theta dx+\sin\theta dy)-\frac{(mr^2-J_0)\tan\phi}{2(mr^2-J_0\sin^2\phi)}\dot{p}d\theta.
\end{aligned}
$$

As for $dH_{\mathcal{H}}$, recall that the constrained Hamiltonian $H_{\mathcal{M}}$ is

$$
H_{\mathcal{M}}=\frac{mr^2}{2(mr^2-J_0)^2}\cot^2\phi(p_\theta-p_\psi)^2+\frac{1}{2J_0}p_\psi^2+\frac{1}{2(mr^2-J_0)}(p_\theta-p_\psi)^2+\frac{1}{4J_1}p_\phi^2.
$$

Notice that $H_{\mathcal{M}}$ is $SE(2)$-invariant and hence $H_{\mathcal{M}}=h_{\overline{\mathcal{M}}}$ where

$$
\begin{aligned}
h_{\overline{\mathcal{M}}} &= \frac{mr^2}{2}\left(\frac{1}{2(mr^2-J_0\sin^2\phi)}p-\frac{\sin^2\phi}{2(mr^2-J_0\sin^2\phi)}p_\psi\right)^2+\frac{1}{2J_0}p_\psi^2\\
&\quad+\frac{mr^2-J_0}{2}\left(\frac{\tan\phi}{2(mr^2-J_0\sin^2\phi)}p-\frac{\sin^2\phi}{mr^2-J_0\sin^2\phi}p_\psi\right)^2+\frac{1}{4J_1}p_\phi^2.
\end{aligned}
$$

Compute $dH_{\mathcal{M}}=dh_{\overline{\mathcal{M}}}$ and we have

$$
\begin{aligned}
&dh_{\overline{\mathcal{M}}}=\\
&\quad\frac{mr^2(p-\sin 2\phi p_\psi)}{2(mr^2-J_0\sin^2\phi)}\left(\frac{1}{2(mr^2-J_0\sin^2\phi)}dp-\frac{\sin 2\phi}{2(mr^2-J_0\sin^2\phi)}dp_\psi\right)\\
&\quad+\frac{mr^2(p-\sin 2\phi p_\psi)}{2(mr^2-J_0\sin^2\phi)}\left(pd\left(\frac{1}{2(mr^2-J_0\sin^2\phi)}\right)-p_\psi d\left(\frac{\sin 2\phi}{2(mr^2-J_0\sin^2\phi)}\right)\right)\\
&\quad+\frac{(mr^2-J_0)(\tan\phi p-2\sin^2\phi p_\psi)}{2(mr^2-J_0\sin^2\phi)}\left(\frac{\tan\phi}{2(mr^2-J_0\sin^2\phi)}dp-\frac{\sin^2\phi}{(mr^2-J_0\sin^2\phi)}dp_\psi\right)\\
&\quad+\frac{(mr^2-J_0)(\tan\phi p-2\sin^2\phi p_\psi)}{2(mr^2-J_0\sin^2\phi)}\left(pd\left(\frac{\tan\phi}{2(mr^2-J_0\sin^2\phi)}\right)-p_\psi d\left(\frac{\sin^2\phi}{(mr^2-J_0\sin^2\phi)}\right)\right)\\
&\quad+\frac{1}{J_0}p_\psi dp_\psi++\frac{1}{2J_1}p_\phi dp_\phi.
\end{aligned}
$$

It is easy to check that $X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}} = dH_{\mathcal{M}}$ is $SE(2)$-invariant, and vanishes on $\mathcal{V} \cap \mathcal{H}$ when restricted to $\mathcal{U}$. Hence both sides push down to $\overline{\mathcal{H}}$. The push down of $X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}}$ is given by

$$
\begin{aligned}
&X_{\overline{\mathcal{H}}} \lrcorner \Omega_{\overline{\mathcal{H}}} = \\
&\left(\frac{J_0}{2mr}\sin(2\phi)\dot{\psi} - \frac{1}{2mr}p\right)\left(\frac{mr\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}dp_\psi - \frac{mr}{2(mr^2 - J_0\sin^2\phi)}dp\right) \\
&+\left(\frac{J_0}{2mr}\sin(2\phi)\dot{\psi} - \frac{1}{2mr}p\right)\left(\frac{mr(mr^2\cos 2\phi + J_0\sin^2\phi)}{(mr^2 - J_0\sin^2\phi)^2}p_\psi d\phi - \frac{mrJ_0\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)^2}pd\phi\right) \\
&+\left(\frac{-J_0}{mr^2}\sin^2(\phi)\dot{\psi} + \frac{\tan\phi}{2mr^2}p\right)\left(\frac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)}dp_\psi + \frac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)}dp\right) \\
&+\left(\frac{-J_0}{mr^2}\sin^2(\phi)\dot{\psi} + \frac{\tan\phi}{2mr^2}p\right)\frac{mr^2(J_0 - mr^2)\sin 2\phi}{(mr^2 - J_0\sin^2\phi)^2}p_\psi d\phi \\
&+\left(\frac{-J_0}{mr^2}\sin^2(\phi)\dot{\psi} + \frac{\tan\phi}{2mr^2}p\right)\frac{(mr^2 - J_0)(mr^2\sec^2\phi + J_0\tan^2\phi\cos 2\phi)}{2(mr^2 - J_0\sin^2\phi)^2}pd\phi \\
&+\dot{\psi}dp_\psi + \dot{\phi}dp_\phi - \dot{p}_\psi d\psi - \dot{p}_\phi d\phi.
\end{aligned}
$$

Equating the terms of $dh_{\overline{\mathcal{H}}} = dh_{\overline{\mathcal{M}}}$ with those of the push down of $X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}}$ gives the remaining reduced Hamilton equations:

$$
\begin{aligned}
\dot{\psi} &= -\frac{\tan\phi}{2(mr^2 - J_0\sin^2\phi)}p + \frac{mr^2}{J_0(mr^2 - J_0\sin^2\phi)}p_\psi && (3.5.10)\\
\dot{\phi} &= \frac{p_\phi}{2J_1} && (3.5.11)\\
\dot{p}_\psi &= 0 && (3.5.12)\\
\dot{p}_\phi &= 0. && (3.5.13)
\end{aligned}
$$

Notice that both the momentum equation (3.5.8) and the above set of reduced equations are independent of the group elements of the symmetry group $SE(2)$. If we add in the set of reconstruction equations (3.5.2), we recover the full dynamics of the system, and in a form that is suitable for control theoretical purposes.

Finding the Reduced Equations on the Lagrangian Side As shown in the proof of Theorem 3.2, we can derive the reduced Lagrange-d'Alembert equations in two ways. Here we will first use the equations (3.4.17).

$$
\frac{d}{dt}\left(\frac{\partial l_c}{\partial \dot{r}^\alpha}\right) - \frac{\partial l_c}{\partial r^\alpha} = -\frac{\partial l}{\partial \xi^b}(B^b_{\alpha\beta}\dot{r}^\beta + F^{bi}p_i), \tag{3.5.14}
$$

where

$$
\begin{aligned}
B^b_{\alpha\beta} &= \frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} - C^b_{ac}A^a_\beta A^c_\alpha \\
F^{bi}_\alpha &= \frac{\partial \Gamma^{bi}}{\partial r^\alpha} - C^b_{ad}A^a_\alpha \Gamma^{di}.
\end{aligned}
$$

From the Lagrangian L, we find the reduced Lagrangian

$$
l(r,\dot{r},\xi) = \frac{1}{2}m((\xi^1)^2 + (\xi^2)^2) + \frac{1}{2}mr^2(\xi^3)^2 + \frac{1}{2}J_0\dot{\psi}^2 + +J_0\dot{\psi}(\xi^3) + J_1\dot{\phi}^2, \tag{3.5.15}
$$

where $\xi = g^{-1}\dot{g}$. After plugging in the constraints (3.5.2), we have the constrained reduced Lagrangian

$$
l_c(r,\dot{r},p) = -\frac{J_0}{2mr^2}\sin^2\phi\dot{\psi}^2 + \frac{1}{8mr^2}\sec^2\phi p^2 + \frac{1}{2}J_0\dot{\psi}^2 + +J_1\dot{\phi}^2. \tag{3.5.16}
$$

Let us find all the ingredients of the above equations:

$$\begin{aligned}\frac{\partial l}{\partial \xi^1} &= m\xi^1 = m\left(\frac{J_0}{2mr}\sin 2\phi\dot{\psi} - \frac{1}{2mr}p\right)\\ \frac{\partial l}{\partial \xi^2} &= m\xi^2 = 0\\ \frac{\partial l}{\partial \xi^3} &= mr^2\left(-\frac{J_0}{mr^2}\sin^2\phi\dot{\psi} + \frac{\tan\phi}{2mr^2}p\right) + J_0\dot{\psi};\end{aligned}$$

since $\frac{\partial l}{\partial \xi^2} = 0$, we do not need to compute $B^2_{\alpha\beta}$ and F^2_α (notice that $i = 1$). Also it is straightforward to find

$$\begin{aligned}B^1_{12} &= \frac{\partial}{\partial\phi}\left(-\frac{J_0}{2mr}\sin 2\phi\right) = -\frac{J_0}{mr}\cos 2\phi\\ B^3_{12} &= \frac{\partial}{\partial\phi}\left(\frac{J_0}{mr}\sin^2\phi\right) = \frac{J_0}{mr}\sin 2\phi\\ F^3_2 &= \frac{\partial}{\partial\phi}\left(\frac{\tan\phi}{2mr^2}\right) = \frac{\sec^2\phi}{2mr^2},\end{aligned}$$

and $F^1_1 = F^3_1 = F^1_2 = 0$. Substituting into (3.5.14), we get the reduced equations after some computations

$$\left(1 - \frac{J_0}{mr^2}\sin^2\phi\right)\ddot{\psi} = \frac{J_0}{2mr^2}\sin 2\phi\dot{\psi}\dot{\phi} - \frac{J_0}{2mr^2}\dot{\phi}p \tag{3.5.17}$$

$$J_1\ddot{\phi} = 0 \tag{3.5.18}$$

It is easy to check that these two equationis are equivalent to the set of reduced equations (3.5.10)-(3.5.13) on the Hamiltonian side through the constrained Lengendre transformation $\mathbb{F}L|\mathcal{D}$.

Next we will find the reduced equations use the equations (3.4.25)

$$\frac{d}{dt}\left(\frac{\partial l_c}{\partial \dot{r}^\alpha}\right) - \frac{\partial l_c}{\partial r^\alpha} = -(\mathcal{K}^{jl}_\alpha p_j p_l + \mathcal{K}^j_{\alpha\beta}\dot{r}^\beta p_j + \mathcal{K}_{\alpha\beta\delta}\dot{r}^\beta\dot{r}^\delta) \tag{3.5.19}$$

where

$$\begin{aligned}\mathcal{K}^{jl}_\alpha &= \frac{\partial I^{jl}}{\partial r^\alpha} - C^j_{bh}A^b_\alpha I^{hl} + \gamma^j_{h\alpha}I^{hl}\\ \mathcal{K}^j_{\alpha\beta} &= \lambda_{a'\beta}(-C^{a'}_{bh}A^b_\alpha I^{hj} + \gamma^{a'}_{h\alpha}I^{hj}) + B^j_{\alpha\beta}\\ \mathcal{K}_{\alpha\beta\delta} &= \lambda_{a'\delta}B^{a'}_{\alpha\beta}.\end{aligned}$$

Here

$$B^b_{\alpha\beta} = \frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} - C^b_{ac}A^a_\beta A^c_\alpha + A^c_\alpha\gamma^b_{c\beta} - A^c_\beta\gamma^b_{c\alpha}.$$

First we need to construct the orthogonal body frame. Recall that

$$(e_1(g,r))_Q = g^a_d e^d_1 \partial_{g^a} = -2r\cos^2\phi\cos\theta\partial_x - 2r\cos^2\phi\sin\theta\partial_y + \sin 2\phi\partial_\theta.$$

Hence

$$e_1 = -2r\cos^2\phi e_x + \sin 2\phi e_\theta,$$

where e_x, e_y, e_θ are the generators of $\partial_x, \partial_y, \partial_\theta$. Using the kinetic energy metric, we find

$$\begin{aligned}e_2 &= -\frac{1}{m}\sin\phi e_x + \frac{1}{m}\cos\phi e_y - \frac{1}{mr}\cos\phi e_\theta\\ e_3 &= \frac{1}{m}\sin\phi e_x + \frac{1}{m}\cos\phi e_y + \frac{1}{mr}\cos\phi e_\theta\end{aligned}$$

Recall that we only need e_1 to be orthogonal to e_2 and e_3.

Let η^b be the components of ξ in the new basis, i.e., $\xi = \xi^1 e_x + \xi^2 2_y + \xi^3 e_\theta = \eta^a e_a$, then

$$\begin{aligned}
\xi^1 &= -2r\cos^2\phi\eta^1 - \frac{1}{m}\sin\phi\eta^2 + \frac{1}{m}\sin\phi\eta^3 \\
\xi^2 &= \frac{1}{m}\cos\phi\eta^2 + \frac{1}{m}\cos\phi\eta^3 \\
\xi^1 &= \sin 2\phi\eta^1 - \frac{1}{mr}\cos\phi\eta^2 + \frac{1}{mr}\cos\phi\eta^3,
\end{aligned}$$

and $\bar{l}(r, \dot{r}, \eta^a) = l(r, \dot{r}, T_a^b \eta^a)$ where T_b^a is defined as above by $\xi^b = T_a^b \eta^a$.

Notice that in the new basis, the constraints (3.5.2) become

$$\begin{bmatrix} \eta^1 \\ \eta^2 \\ \eta^3 \end{bmatrix} = - \begin{bmatrix} \frac{J_0}{2mr^2}\tan\phi\dot{\psi} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{1}{4mr^2}\sec^2\phi p \\ 0 \\ 0 \end{bmatrix} \tag{3.5.20}$$

but the constrained reduced equation $\bar{l}_c(r, \dot{r}, p)$ remains the same and is equal to $l_c(r, \dot{r}, p)$.

Let us find all the ingredients of equations (3.5.19). After finding from (3.5.20) that

$$\begin{aligned}
A_1^1 &= \frac{J_0}{2mr^2}\tan\phi \\
I^{11} &= \frac{1}{4mr^2}\sec^2\phi
\end{aligned}$$

and the rest of A_α^b equal to zero (which is not true in general), it is straightforward to calculate

$$\begin{aligned}
\mathcal{K}_1^{11} &= 0 \\
\mathcal{K}_2^{11} &= \frac{1}{4mr^2}\sec^2\phi\tan\phi \\
\mathcal{K}_{11}^1 &= 0 \\
\mathcal{K}_{12}^1 &= \frac{J_0}{2mr^2} \\
\mathcal{K}_{21}^1 &= 0 \\
\mathcal{K}_{22}^1 &= 0 \\
\mathcal{K}_{121} &= \frac{J_0^2}{2mr^2}\sin 2\phi \\
\mathcal{K}_{122} &= 0
\end{aligned}$$

After substituting into (3.5.19) we get the same reduced equations as (3.5.17) and (3.5.18).

3.6 Example: The Bicycle

Control of the bicycle is a rich problem offering a number of considerable challenges of current research interest in the area of mechanical and robotic control. The bicycle is an underactuated system, subject to nonholonomic contact constraints associated with the rolling constraints on the front and rear wheels. It is unstable (except under certain combinations of fork geometry and speed) when not controlled. It is also, when considered to traverse flat ground, a system subject to symmetries; its Lagrangian and constraints are invariant with respect to translations and rotations in the ground plane.

Here a simplified bicycle model will be considered. The wheels of the bicycle are considered to have negligible inertia moments, mass, radii, and width, and roll without side or longitudinal slip.

The vehicle is assumed to have a fixed steering axis that is perpendicular to the flat ground when the bicycle is upright. For simplicity we concern ourselves with a point mass bicycle. The rigid frame of the bicycle will be assumed to be symmetric about a plane containing the rear wheel.

Consider a ground fixed inertial reference frame with x and y axis in the ground plane and z-axis perpendicular to the ground plane in the direction opposite to gravity. The interseciton of the vehicle's plane of symmetry with the ground plane forms a contact line. The contact line is rotated about the z-direction by a yaw angle θ. The contact line is considered directed, with its positive direction from the rear to the front of the vehicle. The yaw angle θ is zero when the contact line is in the x-direction. The angle that the bicycle's plane of symmetry makes with the vertical direction is the roll angle $\psi \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Front and rear wheel contacts are constrained to have velocities parallel to the lines of intersection of their respective wheel planes and the ground plane, but free to turn about an axis through the wheel/ground contact and parallel to the z-axis. Let $\sigma \in (-\frac{\pi}{2}, \frac{\pi}{2})$ be the steering angle between the front wheel plane/ground plane intersection and the contact line. With σ we associate a moment of inertia J which depends both on ψ and σ. We will parametrize the steering angle by $\phi := \tan\sigma/b$. For more details, see Getz and Marsden [1995] and Getz [1996]. See figure 3.6.1.

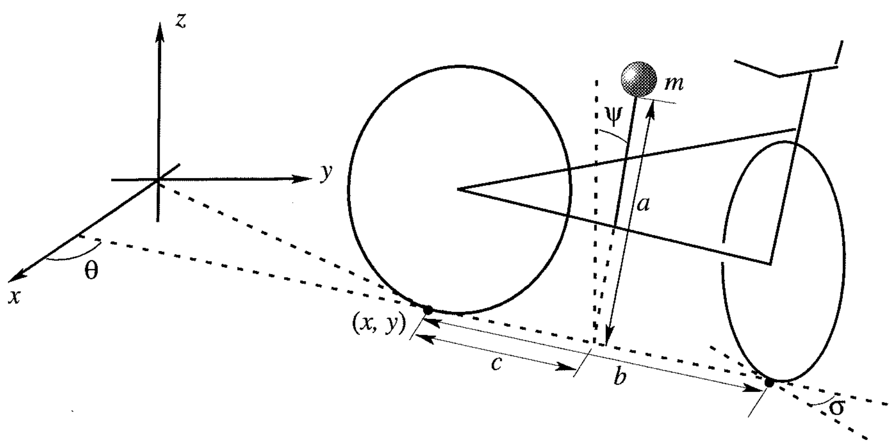

Figure 3.6.1: Notation for the bike.

The configuration space is $Q = SE(2) \times S^1 \times S^1$ and the Lagrangian $L : TQ \to \mathbb{R}$ is the total kinetic energy minus potential energy of the system and is given by

$$\begin{aligned} L &= -mga\cos\psi + \frac{1}{2}J(\psi,\phi)\dot{\phi}^2 \\ &\quad + \frac{m}{2}\left((\cos\theta\dot{x} + \sin\theta\dot{y} + a\sin\psi\dot{\theta})^2 + (-\sin\theta\dot{x} + \cos\theta\dot{y} - a\cos\psi\dot{\psi} + c\dot{\theta})^2 + (-a\sin\psi\dot{\psi})^2\right) \end{aligned}$$

where m is the mass of the bicycle, considered for simplicity to be a point mass, and $J(\psi,\phi)$ is the moment of inertia associated with the steering action. The nonholonomic constraints associated with the front and rear wheels, assumed to roll without slipping, are expressed by

$$\begin{aligned} \dot{\theta} - \phi(\cos\theta\dot{x} + \sin\theta\dot{y}) &= 0 \\ -\sin\theta\dot{x} + \cos\theta\dot{y} &= 0. \end{aligned}$$

Clearly both the Lagrangian and the constraints are invariant under the $SE(2)$ action.

Notice that the Legendre transform $\mathbb{F}L$ is singular but by the remark following Theorem 3.2 the Hamiltonian procedure still works because the constrained Legendre transform $\mathbb{F}L|\mathcal{D}$ is invertible.

The Constraint Submanifold The constraints above give rise to the constraint one forms

$$\begin{aligned}\omega_1(q) &= d\theta - \phi\cos\theta dx - \phi\sin\theta dy\\ \omega_2(q) &= -\sin\theta dx + \cos\theta dy\end{aligned}$$

which determine the kinematic distribution $\mathcal{D}_q$:

$$\mathcal{D}_q = \mathrm{span}\{\partial_\psi, \partial_\phi, \cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta\}.$$

The tangent space to the orbits of the $SE(2)$ action is given by

$$T_q(\mathrm{Orb}(q)) = \mathrm{span}\{\partial_x, \partial_y, \partial_\theta\},$$

and the intersection between the tangent space to the group orbits and the constraint distribution is thus given by

$$\mathcal{S}_q = \mathcal{D}_q \cap T_q(\mathrm{Orb}(q)) = \mathrm{span}\{\cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta\}.$$

The momentum can be constructed by choosing a section of $\mathcal{S} = \mathcal{D}\cap T\mathrm{Orb}$ regarded as a bundle over Q. Since $\mathcal{D}_q \cap T_q\mathrm{Orb}(q)$ is one-dimensional, the section can be chosen to be

$$\xi_Q^q = \cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta,$$

which is invariant under the action of $SE(2)$ on Q. The nonholonomic momentum map is thus given by

$$\begin{aligned}p &= \frac{\partial L}{\partial \dot{q}^i}(\xi_Q^q)^i\\ &= m(\dot{x} + a\sin\psi\cos\theta\dot{\theta} + a\cos\psi\sin\theta\dot{\psi} - c\sin\theta\dot{\theta})\cos\theta\\ &\quad +m(\dot{y} + a\sin\psi\sin\theta\dot{\theta} - a\cos\psi\cos\theta\dot{\psi} + c\cos\theta\dot{\theta})\sin\theta\\ &\quad +m(\cos\theta\dot{x} + \sin\theta\dot{y} + a\sin\psi\dot{\theta})a\phi\sin\psi\\ &\quad +m(-\sin\theta\dot{x} + \cos\theta\dot{y} - a\cos\psi\dot{\psi} + c\dot{\theta})c\phi.\end{aligned}$$

The kinematic constraints plus the momentum are given by

$$\begin{aligned}0 &= \xi^3 - \phi\xi^1\\ 0 &= \xi^2\\ p &= m(\xi^1 + a\sin\psi\xi^3) + ma\phi\sin\psi(\xi^1 + a\sin\psi\xi^3)\\ &\quad m\phi(c\xi^2 - ca\cos\psi\dot{\psi} + c^2\xi^3)\end{aligned}$$

where

$$\begin{aligned}\xi^1 &= \cos\theta\dot{x} + \sin\theta\dot{y}\\ \xi^2 &= -\sin\theta\dot{x} + \cos\theta\dot{y}\\ \xi^3 &= \dot{\theta}\end{aligned}$$

Adding, subtracting, and scaling these equations, we can write

$$\begin{bmatrix}\xi^1\\ \xi^2\\ \xi^3\end{bmatrix} + \begin{bmatrix}-\dfrac{ca\phi\cos\psi}{K}\dot{\psi}\\ 0\\ -\dfrac{ca\phi^2\cos\psi}{K}\dot{\psi}\end{bmatrix} = \begin{bmatrix}\dfrac{1}{mK}p\\ 0\\ \dfrac{\phi}{mK}p\end{bmatrix} \tag{3.6.1}$$

where

$$K = (1 + a\phi \sin\psi)^2 + c^2\phi^2. \tag{3.6.2}$$

These equations have the form

$$g^{-1}\dot{g} + A(r)\dot{r} = \Gamma(r)p.$$

Next find the Legendre transform $\mathbb{F}L$ and restrict it to the constraint submanifold $\mathcal{D} \subset TQ$, we get

$$\begin{aligned}
p_x &= m(1 + a\phi\sin\psi)\xi^1\cos\theta - m(c\phi\xi^1 - a\cos\psi\dot{\psi})\sin\theta \\
p_y &= m(1 + a\phi\sin\psi)\xi^1\sin\theta + m(c\phi\xi^1 - a\cos\psi\dot{\psi})\cos\theta \\
p_\theta &= ma\sin\psi(1 + a\phi\sin\psi)\xi^1 + m(c^2\phi\xi^1 - ca\cos\psi\dot{\psi}) \\
p_\psi &= ma^2\dot{\psi} - mac\cos\psi\phi\xi^1 \\
p_\phi &= J(\psi, \phi)\dot{\phi}.
\end{aligned}$$

After applying the constrained Legendre transformation $\mathbb{F}L|\mathcal{D}$ and its inverse to the constraint equations (3.6.1), we have

$$\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \begin{bmatrix} -\dfrac{c\phi\cos\psi(1 + a\phi\sin\psi)}{F}\dfrac{p_\psi}{a} \\ \dfrac{(1 + a\phi\sin\psi)^2\cos\psi}{F}\dfrac{p_\psi}{a} \\ \dfrac{c\cos\psi(1 + a\phi\sin\psi)}{F}\dfrac{p_\psi}{a} \end{bmatrix} = \begin{bmatrix} \dfrac{1 + a\phi\sin\psi}{F}p \\ \dfrac{c\phi\sin^2\psi}{F}p \\ \dfrac{(1 + a\phi\sin\psi)a\sin\psi + c^2\phi\sin^2\psi}{F}p \end{bmatrix}, \tag{3.6.3}$$

where

$$\begin{aligned}
\mu_1 &= \cos\theta p_x + \sin\theta p_y \\
\mu_2 &= -\sin\theta p_x + \cos\theta p_y \\
\mu_3 &= p_\theta
\end{aligned}$$

and

$$F = (1 + a\phi\sin\psi)^2 + c^2\phi^2\sin^2\psi \tag{3.6.4}$$

$$p = p_x\cos\theta + p_y\sin\theta + p_\theta\phi. \tag{3.6.5}$$

Therefore, the constraint submanifold $\mathcal{M} \subset T^*Q$ is defined by

$$\begin{aligned}
p_x &= \mu_1\cos\theta - \mu_2\sin\theta \\
p_y &= \mu_1\sin\theta + \mu_2\cos\theta \\
p_\theta &= \mu_3.
\end{aligned}$$

It is a submanifold in T^*Q and we can use $(x, y, \theta, \psi, \phi, p_\psi, p_\phi, p)$ as its induced local coordinates.

The Distributions $\mathcal{H}, \mathcal{V} \cap \mathcal{H}$ and $\mathcal{U}$. Using the induced coordinates, the distribution $\mathcal{H}$ on $\mathcal{M}$ is

$$\mathcal{H} = \mathrm{span}\{\cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta, \partial_\psi, \partial_\phi, \partial_{p_\psi}, \partial_{p_\phi}, \partial_p\} \tag{3.6.6}$$

and the subdistribution $\mathcal{V} \cap \mathcal{H}$ is

$$\mathcal{V} \cap \mathcal{H} = \mathrm{span}\{\cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta\}. \tag{3.6.7}$$

Notice that in the case of the bicycle, the constraints are independent of the velocities of the shape variables and hence the simplified procedure employed in the snakeboard is also used here.

As for the subdistribution $\mathcal{U}$, we first calculate the two form $\Omega_{\mathcal{M}}$. After pulling back the canonical two-form of T^*Q to $\mathcal{M}$, we have

$$\begin{aligned}
\Omega_{\mathcal{M}} &= dx \wedge dp_x + dy \wedge dp_y + d\theta \wedge dp_\theta + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi \\
&= (\cos\theta dx + \sin\theta dy) \wedge d\mu_1 + \mu_1(-\sin\theta dx + \cos\theta dy) \wedge d\theta \\
&\quad +(-\sin\theta dx + \cos\theta dy) \wedge d\mu_2 - \mu_2(\cos\theta dx + \sin\theta dy) \wedge d\theta \\
&\quad +d\theta \wedge d\mu_3 + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi
\end{aligned}$$

Since $\mathcal{U} = (\mathcal{V} \cap \mathcal{H})^\perp = \ker\{(\mathcal{V} \cap \mathcal{H}) \lrcorner \Omega_{\mathcal{H}}\}$, we need to calculate $(\mathcal{V} \cap \mathcal{H}) \lrcorner \Omega_{\mathcal{M}}$, and restrict it to $\mathcal{H}$:

$$\begin{aligned}
(\mathcal{V} \cap \mathcal{H}) \lrcorner \Omega_{\mathcal{H}} &= d\mu_1 - \mu_1\phi(-\sin\theta dx + \cos\theta dy) \\
&\quad -\mu_2 d\theta + \mu_2\phi(\cos\theta dx + \sin\theta dy) + \phi d\mu_3 \\
&= d\mu_1 + \phi d\mu_3 \\
&= dp + \frac{c\cos\psi(1 + a\phi\sin\psi)}{F}\frac{p_\psi}{a}d\phi - \frac{a\sin\psi(1 + a\phi\sin\psi) + c^2\phi\sin^2\psi}{F}p d\phi.
\end{aligned}$$

Hence,

$$\mathcal{U} = \ker\left\{dp + \frac{c\cos\psi(1 + a\phi\sin\psi)}{F}\frac{p_\psi}{a}d\phi - \frac{a\sin\psi(1 + a\phi\sin\psi) + c^2\phi\sin^2\psi}{F}p d\phi\right\}. \tag{3.6.8}$$

The Reconstruction and Momentum Equations A vector field $X_{\mathcal{U}}$ taking values in $\mathcal{U}$ must be of the form

$$X_{\mathcal{U}} = \dot{x}\partial_x + \dot{y}\partial_y + \dot{\theta}\partial_\theta + \dot{\psi}\partial_\psi + \dot{\phi}\partial_\phi + \dot{p}_\psi\partial_{p_\psi} + \dot{p}_\phi\partial_{p_\phi} + \dot{p}\partial_p \tag{3.6.9}$$

where

$$\begin{aligned}
\dot{x} &= \xi^1\cos\theta - \xi^2\sin\theta = \left(\frac{ca\phi\cos\psi}{K}\dot{\psi} + \frac{1}{mK}p\right)\cos\theta \\
\dot{y} &= \xi^1\sin\theta + \xi^2\cos\theta = \left(\frac{ca\phi\cos\psi}{K}\dot{\psi} + \frac{1}{mK}p\right)\sin\theta \\
\dot{\theta} &= \phi\xi^1 = \left(\frac{ca\phi^2\cos\psi}{K}\dot{\psi} + \frac{\phi}{mK}p\right)
\end{aligned}$$

and

$$\dot{p} = -\frac{c\cos\psi(1 + a\phi\sin\psi)}{F}\frac{p_\psi}{a}\dot{\phi} + \frac{a\sin\psi(1 + a\phi\sin\psi) + c^2\phi\sin^2\psi}{F}p\dot{\phi}. \tag{3.6.10}$$

The equations for $\dot{x}$, $\dot{y}$ and $\dot{\theta}$ are the same reconstruction equations as equations (3.6.1) and the last one for $\dot{p}$ is the momentum eqution on the Hamiltonian side. Similar to the example of the snakeboard, the momentum p equals the angular momentum of the sysem about a fixed point P that can be determined in the same way as in the case of the snakeboard. Notice also that the last equation can be written simply as $\dot{p} = \mu_3\dot{\phi}$.

The Reduced Hamilton Equations. To find the remaining reduced equations, we need to compute

$$X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}} = dH_{\mathcal{M}}, \tag{3.6.11}$$

restrict it to the subdistribution $\mathcal{U}$ and then push it down to the reduced constraint submanifold $\overline{\mathcal{M}}$. Let us first compute $X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}}$

$$X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}} =$$

$$
\begin{aligned}
&(\cos\theta\dot{x}+\sin\theta\dot{y})d\mu_1+\mu_1(-\sin\theta\dot{x}+\cos\theta\dot{y})d\theta-\mu_1\dot{\theta}(-\sin\theta dx+\cos\theta dy)\\
&+(-\sin\theta\dot{x}+\cos\theta\dot{y})d\mu_2-\mu_2(\cos\theta\dot{x}+\sin\theta\dot{y})d\theta+\mu_2\dot{\theta}(\cos\theta dx+\sin\theta dy)\\
&+\dot{\theta}d\mu_3+\dot{\psi}dp_\psi+\dot{\phi}dp_\phi-\dot{p}_\psi d\psi-\dot{p}_\phi d\phi\\
&-((\dot{\psi}\partial_\psi+\dot{\phi}\partial_\phi+\dot{p}_\psi\partial_{p_\psi}+\dot{p}_\phi\partial_{p_\phi}+\dot{p}\partial_p)\lrcorner\, d\mu_1)(\cos\theta dx+\sin\theta dy)\\
&-((\dot{\psi}\partial_\psi+\dot{\phi}\partial_\phi+\dot{p}_\psi\partial_{p_\psi}+\dot{p}_\phi\partial_{p_\phi}+\dot{p}\partial_p)\lrcorner\, d\mu_2)(-\sin\theta dx+\cos\theta dy)
\end{aligned}
$$

As for $dH_{\mathcal{H}}$, we can find the constrained Hamiltonian $H_{\mathcal{M}}$ via the constrained Legendre transform and have

$$
\begin{aligned}
H_{\mathcal{M}} &= mga\cos\psi+\frac{1}{2J}p_\phi^2+\\
&\quad \frac{1}{2m}\left(\mu_1^2+\mu_2^2+\left(\frac{K\sin\psi}{F}\frac{p_\psi}{a}+\frac{c\phi\sin\psi\cos\psi}{F}p\right)^2\right).
\end{aligned}
$$

Notice that $H_{\mathcal{M}}$ is $SE(2)$-invariant and hence $H_{\mathcal{M}}=h_{\overline{\mathcal{M}}}$. Compute $dH_{\mathcal{M}}=dh_{\overline{\mathcal{M}}}$ and we have

$$
\begin{aligned}
&dh_{\overline{\mathcal{M}}}=\\
&\quad -mga\sin\psi d\psi+\frac{1}{J}p_\phi dp_\phi-\frac{1}{2J^2}p_\phi^2\left(\frac{\partial J}{\partial\psi}d\psi+\frac{\partial J}{\partial\phi}d\phi\right)\\
&\quad +\frac{1}{m}\left(\mu_1d\mu_1+\mu_2d\mu_2+\left(\frac{K\sin\psi}{F}\frac{p_\psi}{a}+\frac{c\phi\sin\psi\cos\psi}{F}p\right)d\left(\frac{K\sin\psi}{F}\frac{p_\psi}{a}+\frac{c\phi\sin\psi\cos\psi}{F}p\right)\right).
\end{aligned}
$$

It can be checked that $X_{\mathcal{H}}\lrcorner\,\Omega_{\mathcal{M}}=dH_{\mathcal{M}}$ is $SE(2)$-invariant, and vanishes on $\mathcal{V}\cap\mathcal{H}$ when restricted to $\mathcal{U}$. Hence both sides push down to $\overline{\mathcal{H}}$. The push down of $X_{\mathcal{H}}\lrcorner\,\Omega_{\mathcal{M}}$ is given by

$$
\begin{aligned}
X_{\overline{\mathcal{H}}}\lrcorner\,\Omega_{\overline{\mathcal{H}}} &= (\cos\theta\dot{x}+\sin\theta\dot{y})d\mu_1+\dot{\theta}d\mu_3+\dot{\psi}dp_\psi+\dot{\phi}dp_\phi-\dot{p}_\psi d\psi-\dot{p}_\phi d\phi\\
&= \xi^1d\mu_1+\xi^3d\mu_3+\dot{\psi}dp_\psi+\dot{\phi}dp_\phi-\dot{p}_\psi d\psi-\dot{p}_\phi d\phi
\end{aligned}
$$

Equating the terms of $dh_{\overline{\mathcal{H}}}=dh_{\overline{\mathcal{M}}}$ with those of the push down of $X_{\mathcal{H}}\lrcorner\,\Omega_{\mathcal{M}}$ gives the remaining reduced Hamilton equations:

$$
\dot{\psi} = \frac{1}{ma}\left(\frac{K}{F}\frac{p_\psi}{a}+\frac{c\phi\cos\psi}{F}p\right) \tag{3.6.12}
$$

$$
\dot{\phi} = \frac{p_\phi}{J} \tag{3.6.13}
$$

$$
\dot{p}_\psi = mga\sin\psi+\frac{1}{2J^2}p_\phi^2\frac{\partial J}{\partial\psi}+m(1+a\phi\sin\psi)a\phi\cos\psi(\xi^1)^2+mca\phi\sin\psi\xi^1\dot{\psi} \tag{3.6.14}
$$

$$
\dot{p}_\phi = \frac{1}{2J^2}\frac{\partial J}{\partial\phi}p_\phi^2, \tag{3.6.15}
$$

where

$$
\xi^1=\frac{c\phi\cos\psi}{K}\dot{\psi}+\frac{1}{mK}p=\frac{c\phi\cos\psi}{mF}\frac{p_\psi}{a}+\frac{1}{mF}p
$$

as defined earlier in (3.6.1). The first two equations are nothing but the inverse of the constrained Legendre transform. Notice that both the momentum equation (3.6.10) and the above set of reduced equations are independent of the group elements of the symmetry group $SE(2)$. If we add in the set of reconstruction equations (3.6.1), we recover the full dynamics of the system, and in a form that is suitable for control theoretical purposes.

3.7 Example: A Nonholonomically Constrained Particle

In [BS], the example of a nonholonomically constrained particle has been used to illustrate its theory. Here, we would like to modify this example slightly in order to show concretely what need to be done to find the reduced equations of motion if the constraints involve also the velocities of the shape variables.

Consider a particle with the Lagrangian

$$L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \tag{3.7.1}$$

and the nonholonomic constraint

$$\dot{z} = y\dot{x} + \dot{y}. \tag{3.7.2}$$

The constraint and the Lagrangian are invariant under the $\mathbb{R}^2$ action on $\mathbb{R}^3$ given by

$$(x, y, z) \mapsto (x + \lambda, y, z + \mu).$$

Notice that in the original example used in [BS], the constraint does not involve the $\dot{y}$-term and hence it also satisfies the special condition that the constraints are independent of the velocities of the shape variables. But the slight modification changes all these.

The Constraint Submanifold The constraint above gives rise to the constraint one form

$$\omega(q) = dz - ydx - dy. \tag{3.7.3}$$

The tangent space to the orbits of this group action is given by

$$T_q(\mathrm{Orb}(q)) = \mathrm{span}\{\partial_x, \partial_z\},$$

and the intersection between the tangent space to the group orbits and the constraint distribution is thus given by

$$\mathcal{S}_q = \mathcal{D}_q \cap T_q(\mathrm{Orb}(q)) = \mathrm{span}\{\partial_x + y\partial_z\}.$$

The momentum can be constructed by choosing a section of $\mathcal{S} = \mathcal{D} \cap T\mathrm{Orb}$ regarded as a bundle over Q. Since $\mathcal{D}_q \cap T_q\mathrm{Orb}(q)$ is one-dimensional, the section can be chosen to be

$$\xi^q_Q = \partial_x + y\partial_z.$$

The nonholonomic momentum map is thus given by

$$p = \frac{\partial L}{\partial \dot{q}^i}(\xi^q_Q)^i = \dot{x} + y\dot{z}. \tag{3.7.4}$$

The kinematic constraint plus the momentum are

$$\begin{aligned} -y\dot{x} + \dot{z} &= \dot{y} \\ \dot{x} + y\dot{z} &= p. \end{aligned}$$

Solving for $\dot{x}$ and $\dot{z}$, we get

$$\dot{x} = -\frac{y}{1+y^2}\dot{y} + \frac{1}{1+y^2}p \tag{3.7.5}$$

$$\dot{z} = \frac{1}{1+y^2}\dot{y} + \frac{y}{1+y^2}p. \tag{3.7.6}$$

After applying the constrained Legendre transform, we find that the constraint submanifold $\mathcal{M} \subset T^*Q$ is defined by

$$p_x = -\frac{y}{1+y^2}p_y + \frac{1}{1+y^2}p \tag{3.7.7}$$

$$p_z = \frac{1}{1+y^2}p_y + \frac{y}{1+y^2}p. \tag{3.7.8}$$

It is a submanifold in T^*Q and we can use (x, y, z, p_y, p) as its induced local coordinates.

The Distributions $\mathcal{H}, \mathcal{V} \cap \mathcal{H}$ and $\mathcal{U}$. With the induced coordinates, the distribution $\mathcal{H}$ on $\mathcal{M}$ is

$$\mathcal{H} = \text{span}\left\{\partial_x + y\partial_z, -\frac{y}{1+y^2}\partial_x + \frac{1}{1+y^2}\partial_z + \partial_y, \partial_{p_y}, \partial_p\right\} \tag{3.7.9}$$

Notice that we are using $-g^a_b A^b_\alpha \partial_{g^a} + \partial_{r^\alpha}$, i.e, $-\frac{y}{1+y^2}\partial_x + \frac{1}{1+y^2}\partial_z + \partial_y$ instead of ∂_y. In fact, ∂_y does not even lie in the distribution $\mathcal{H}$.

The subdistribution $\mathcal{V} \cap \mathcal{H}$ is

$$\mathcal{V} \cap \mathcal{H} = \text{span}\{\partial_x + y\partial_x\}. \tag{3.7.10}$$

As for the subdistribution $\mathcal{U}$, we first calculate the two form $\Omega_\mathcal{M}$. After pulling back the canonical two-form of T^*Q to $\mathcal{M}$, we have

$$\begin{aligned}
\Omega_\mathcal{M} &= dx \wedge dp_x + dy \wedge dp_y + dz \wedge dp_z \\
&= dx \wedge d\left(-\frac{y}{1+y^2}p_y + \frac{1}{1+y^2}p\right) \\
&\quad + dy \wedge dp_y + dz \wedge d\left(\frac{1}{1+y^2}p_y + \frac{y}{1+y^2}p\right).
\end{aligned}$$

Since $\mathcal{U} = (\mathcal{V} \cap \mathcal{H})^\perp = \ker\{(\mathcal{V} \cap \mathcal{H}) \lrcorner \Omega_\mathcal{H}\}$, we need to calculate $(\mathcal{V} \cap \mathcal{H}) \lrcorner \Omega_\mathcal{M}$, and restrict it to $\mathcal{H}$:

$$\begin{aligned}
(\partial_x + y\partial_z) \lrcorner \Omega_\mathcal{H} &= -\frac{y}{1+y^2}dp_y - \frac{1-y^2}{(1+y^2)^2}p_y dy + \frac{1}{1+y^2}dp - \frac{2y}{(1+y^2)^2}pdy \\
&\quad + \frac{y}{1+y^2}dp_y - \frac{2y^2}{(1+y^2)^2}p_y dy + \frac{y^2}{1+y^2}dp + \frac{y(1-y^2)}{(1+y^2)^2}pdy \\
&= dp - \frac{1}{1+y^2}p_y dy - \frac{y}{1+y^2}pdy
\end{aligned}$$

Hence,

$$\mathcal{U} = \ker\left\{dp - \frac{1}{1+y^2}p_y dy - \frac{y}{1+y^2}pdy\right\}. \tag{3.7.11}$$

The Reconstruction and Momentum Equations A vector field $X_\mathcal{U}$ taking values in $\mathcal{U}$ must be of the form

$$X_\mathcal{U} = \dot{x}\partial_x + \dot{y}\partial_y + \dot{z}\partial_z + \dot{p}_y\partial_{p_y} + \dot{p}\partial_p \tag{3.7.12}$$

where

$$\dot{x} = -\frac{y}{1+y^2}\dot{y} + \frac{1}{1+y^2}p \tag{3.7.13}$$

$$\dot{z} = \frac{1}{1+y^2}\dot{y} + \frac{y}{1+y^2}p. \tag{3.7.14}$$

and

$$\dot{p} - \frac{1}{1+y^2}p_y\dot{y} - \frac{y}{1+y^2}p\dot{y} = 0 \tag{3.7.15}$$

The first set are the reconstruction equations and the last one is the momentum equation on the Hamiltonian side.

The Reduced Hamilton Equations. To find the remaining reduced equations, we need to compute

$$X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}} = dH_{\mathcal{M}}, \tag{3.7.16}$$

restrict it to the subdistribution $\mathcal{U}$ and then push it down to the reduced constraint submanifold $\overline{\mathcal{M}}$. Let us first compute $X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}}$

$$\begin{aligned}
X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}} &= \left(\frac{1-y^2}{(1+y^2)^2}\dot{y}p_y + \frac{2y(1-y^2)}{(1+y^2)^2}\dot{y}p + \frac{y}{1+y^2}\dot{p}_y - \frac{1}{1+y^2}\dot{p}\right)dx \\
&+ \left(\frac{2y}{(1+y^2)^2}\dot{y}p_y - \frac{1-y^2}{(1+y^2)^2}\dot{y}p - \frac{1}{1+y^2}\dot{p}_y - \frac{y}{1+y^2}\dot{p}\right)dz \\
&+ \left(-\frac{1-y^2}{(1+y^2)^2}\dot{x}p_y - \frac{2y}{(1+y^2)^2}\dot{x}p - \frac{2y}{(1+y^2)^2}\dot{z}p_y + \frac{1-y^2}{(1+y^2)^2}\dot{z}p - \dot{p}_y\right)dy \\
&+ \left(-\frac{y}{1+y^2}\dot{x} + \dot{y} + \frac{1}{1+y^2}\dot{z}\right)dp_y + \left(\frac{1}{1+y^2}\dot{x} + \frac{y}{1+y^2}\dot{z}\right)dp
\end{aligned}$$

Notice that in pushing down $X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}}$, we cannot simply just throw away the terms involving dx and dz, instead we have to replace them by $-A^a_\alpha dr^\alpha$, i.e.,by $-\frac{y}{1+y^2}dy$ and $\frac{1}{1+y^2}dy$ respectively, as it has been done in the proof of Theorem 3.2.

As for $dH_{\mathcal{H}}$, we first find the constrained Hamiltonian $H_{\mathcal{M}}$

$$\begin{aligned}
H_{\mathcal{M}} &= \frac{1}{2}\left(\left(-\frac{y}{1+y^2}p_y + \frac{1}{1+y^2}p\right)^2 + p_y^2 + \left(\frac{1}{1+y^2}p_y + \frac{y}{1+y^2}p\right)^2\right) \\
&= \frac{1}{2}\left(\frac{p^2}{1+y^2} + p_y^2 + \frac{p_y^2}{1+y^2}\right)
\end{aligned}$$

Clearly $H_{\mathcal{M}}$ is $\mathbb{R}^2$-invariant and hence $H_{\mathcal{M}} = h_{\overline{\mathcal{M}}}$. Compute $dH_{\mathcal{M}} = dh_{\overline{\mathcal{M}}}$ and we have

$$dh_{\overline{\mathcal{M}}} = \frac{1}{1+y^2}pdp - \frac{y}{(1+y^2)^2}p^2dy + p_ydp_y + \frac{1}{1+y^2}p_ydp_y - \frac{y}{(1+y^2)^2}p_y^2dy$$

Equating the terms of $dh_{\overline{\mathcal{H}}} = dh_{\overline{\mathcal{M}}}$ with those of the push down of $X_{\mathcal{H}} \lrcorner \Omega_{\mathcal{M}}$ gives the remaining reduced Hamilton equations:

$$\dot{y} = p_y \tag{3.7.17}$$

$$\dot{p}_y = -\frac{1}{2+y^2}p_xp_y, \tag{3.7.18}$$

where

$$p_x = -\frac{y}{1+y^2}p_y + \frac{1}{1+y^2}p,$$

as defined earlier in equation (3.7.7).

Conclusions.

In this paper we have analyzed the relation between the Lagrangian and Hamiltonian approaches to problems in nonholonomic mechanics. In the course of doing this, we have clarified each of the pictures. For example, we have shown how the momentum equation first found on the Lagrangian side fits into the Hamiltonian approach. We have also explored the reduced Lagrange-d'Alembert equations in greater detail than was known previously. An example, a simplified model of the bicycle is used to illustrate the ideas.

One aspect we do not address in this paper is the point of view of Poisson geometry and the Dirac theory of constraints. It is known that the obvious Poisson structures for nonholonomic systems do not satisfy the Jacobi identity (this is already mentioned in [BS] and van der Schaft and Maschke [1994]). Thus, any discussion in this direction should take this into account. We hope to address some of these issues in the future. Another item for future work is the extension of the theory of geometric phases (as in Marsden, Montgomery and Ratiu [1990]) to the nonholonomic case.

References

Arnold, V.I. [1988] *Dynamical Systems III.* Encyclopedia of Mathemaics **3**, Springer-Verlag.

Bates, L. and J. Sniatycki [1993] Nonholonomic reduction, *Reports on Math. Phys.* **32**, 99–115.

Bloch, A.M. and P. Crouch [1992] On the dynamics and control of nonholonomic systems on Riemannian Manifolds. *Proceedings of NOLCOS '92*, Bordeaux, 368–372.

Bloch, A.M. and P. Crouch [1994] Nonholonomic control systems on Riemannian manifolds. *SIAM J. on Control* (to appear).

Bloch, A.M. and P. Crouch [1995a] Reduction of Euler Lagrange problems and relation with optimal control problems. *Proc 33rd CDC* (to appear).

Bloch, A.M. and P. Crouch [1995b] Reduction of Euler Lagrange problems for constrained variational problems and relation with optimal control problems. *33rd CDC, to appear.*

Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden, and R. Murray [1996] Nonholonomic mechanical systems with symmetry. *Arch. Rat. Mech. An.* (to appear). CDS Technical Report 94-013, California Institute of Technology, URL http://avalon.caltech.edu/cds/reports/cds94-013.ps.

Getz, N.H. [1994] Control of balance for a nonlinear nonholonomic non-minimum phase model of a bicycle. *ACC Conf.*, Baltimore, June, 1994.

Getz, N.H. & J. E. Marsden [1995] Control for an autonomous bicycle, *International Conference on Robotics and Automation, IEEE*, Nagoya, Japan, May, 1995.

Koiller, J. [1992] Reduction of some classical nonholonomic systems with symmetry. *Arch. Rat. Mech. An.* **118**, 113–148.

Lewis, A. [1996] (preprint, Caltech).

Marsden, J.E. [1992], *Lectures on Mechanics* London Mathematical Society Lecture Note Series. **174**, Cambridge University Press.

Marsden, J.E., R. Montgomery, and T.S. Ratiu [1990] *Reduction, symmetry, and phases in mechanics.* Memoirs AMS **436**.

Marsden, J.E. and T.S. Ratiu [1992] *An Introduction to Mechanics and Symmetry.* Texts in Appl. Math. **17**, Springer-Verlag.

Marsden, J.E. and J. Scheurle [1993a] Lagrangian reduction and the double spherical pendulum, *ZAMP* **44**, 17–43.

Marsden, J.E. and J. Scheurle [1993b] The reduced Euler-Lagrange equations, *Fields Institute Comm.* **1**, 139–164.

Ostrowski, J. [1996] Need to fill in.

van der Schaft, A.J. & B.M. Maschke [1994] On the Hamiltonian formulation of nonholonomic mechanical systems, *Rep. on Math. Phys.* **34**, 225–233.

Vershik, A.M. and V.Ya. Gershkovich [1994] Nonholonomic dynamical systems, geometry of distributions and variational problems, *Dynamical Systems VII, V. Arnold and S.P. Novikov, eds.*, 1–81. Springer-Verlag, New York.

Weber, R.W. [1986] Hamiltonian systems with constraints and their meaning in mechanics. *Arch. Rat. Mech. An.* **91**, 309–335.